Избранные главы
ВЫСШЕЙ МАТЕМАТИКИ
для инженеров и студентов втузов

Л. З. РУМШИСКИЙ

ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ



Л. З. РУМШИСКИЙ

ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

ИЗДАНИЕ ВТОРОЕ, СТЕРЕОТИПНОЕ

Допущено Министерством высшего и среднего специального образования СССР в качестве учебного пособия для высших технических учебных заведений



ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1963

АННОТАЦИЯ

Книга является учебным пособием по курсу теории вероятностей, читаемому в ряде втузов, и соответствует утвержденной программе. Она заполняет имеющийся в нашей литературе пробел между университетскими курсами, слишком трудными для студентов втузов, и популярными книгами, которые содержат не весь необходимый материал. Для понимания книги достаточно знакомства со втузовским курсом математического анализа. Помимо студентов, она может быть полезна инженерам, особенно машиностроительных и радиотехнических специальностей, и экономистам.

ОГЛАВЛЕНИЕ

Предисловие 5
Введение 7

Глава I. **Случайные события и вероятности** 10

§ 1. Случайные события. Относительная частота и вероятность 10
§ 2. Классическое определение вероятности 11
§ 3. Основные свойства вероятностей. Правило сложения вероятностей 14
§ 4. Совмещение случайных событий. Независимые случайные события 19
§ 5. Условные вероятности. Общее правило умножения вероятностей. Формула полной вероятности 23

Глава II. **Случайные величины и распределения вероятностей** 29

§ 6. Дискретные случайные величины 29
§ 7. Распределение вероятностей относительной частоты случайного события 35
§ 8. Непрерывные случайные величины 41
§ 9. Функции от случайных величин 50

Глава III. **Числовые характеристики распределения вероятностей.** 60

§ 10. Осреднение. Математическое ожидание случайной величины 60
§ 11. Центр распределения случайной величины 67
§ 12. Характеристики рассеяния случайной величины. Понятие о моментах распределения 71

Глава IV. **Закон больших чисел** 80

§ 13. О случайных событиях с очень малыми вероятностями 80
§ 14. Теорема Я. Бернулли и устойчивость относительных частот 83
§ 15. Теорема Чебышева 85
§ 16. Устойчивость выборочных средних и метод моментов 89

Глава V. **Предельные теоремы и оценки средних** 98

§ 17. Понятие о характеристических функциях 98
§ 18. Предельная теорема Муавра — Лапласа; оценка относительных частот 102
§ 19. Доверительные оценки средних. Понятие о центральной предельной теореме Ляпунова 107

Глава VI. **Применение теории вероятностей к математической обработке результатов измерений** 116

§ 20. Случайные ошибки измерения, их распределение . . . 116
§ 21. Решение двух основных задач теории ошибок. Оценка истинного значения измеряемой величины и оценка точности прибора в случае прямых равноточных измерений 118

Глава VII. **Линейная корреляция** 133

§ 22. О различных типах зависимостей 133
§ 23. Условные математические ожидания и их свойства . . 135
§ 24. Линейная корреляция 138
§ 25. Коэффициент корреляции 142
§ 26. Наилучшее линейное приближение к функции регрессии 144
§ 27. Анализ линейной корреляции по данным случайной выборки. Оценка значимости коэффициента корреляции . 147

Приложение. Значения интеграла вероятностей 154

ПРЕДИСЛОВИЕ

Настоящее учебное пособие предназначается для высших технических учебных заведений, в которых программой по высшей математике предусмотрено изложение элементов теории вероятностей и математической статистики в объеме до 30 часов. У читателя предполагается знание математического анализа в объеме обычной программы для втузов. Пособие знакомит с основными понятиями и некоторыми методами теории вероятностей, которые в настоящее время необходимы во многих областях техники. Теория случайных процессов и некоторые специальные вопросы, требующие более серьезной математической подготовки, в пособии не освещены. Некоторые более трудные части курса, выходящие за рамки минимальной программы, напечатаны мелким шрифтом и при первом чтении могут быть опущены. Примеры, приводимые в тексте, играют существенную роль в разъяснении основных понятий; рекомендуется разбирать их подробно.

В основу настоящего пособия положен курс лекций, читаемый автором на протяжении десяти лет в Московском энергетическом институте.

При работе над рукописью неоценимую помощь автору оказали советы и замечания А. М. Яглома, которому автор приносит свою глубокую благодарность.

Автор пользуется также возможностью выразить искреннюю благодарность Р. Я. Берри, И. А. Брину, М. И. Вишику, С. А. Стебакову и Р. С. Хасьминскому за ряд полезных указаний.

ВВЕДЕНИЕ

В различных областях техники и производства все чаще приходится иметь дело с массовыми явлениями. Таким массовым явлениям присущи свои особые закономерности. Рассмотрим для примера процесс обработки деталей на станке-автомате. Размеры различных деталей будут колебаться около некоторого установленного значения. Эти колебания носят случайный характер, так что измерения уже обработанных деталей не позволяют точно предсказать величину размера следующей детали. Однако распределение размеров деталей в больших партиях обнаруживает довольно точные закономерности; так, средние арифметические из размеров деталей в разных партиях оказываются приблизительно одинаковыми; отклонения той или иной величины от среднего размера также встречаются в разных партиях почти одинаково часто. Подобное положение можно наблюдать и при многократном взвешивании одного и того же тела на аналитических весах. И здесь отдельные результаты измерений отличаются друг от друга; тем не менее средний результат многих измерений остается практически неизменным; что же касается отклонений от этого среднего результата, то можно точно рассчитать, как часто будут встречаться отклонения той или иной величины. Закономерности такого рода не позволяют, конечно, предсказать результат единичного измерения, но позволяют производить обработку результатов массовых измерений.

Закономерности, относящиеся к различным средним характеристикам, к повторяемости случайных отклонений данной величины и т. п., изучаются специальной математической наукой — теорией вероятностей. Впервые такие закономерности были подмечены при решении задач, связанных с азартными играми, особенно игрой в кости (XVII век). При многократном бросании игральной кости было замечено,

что каждое число очков от 1 до 6 выпадает приблизительно одинаково часто или, как говорят, повторяется с относительной частотой, близкой к $\frac{1}{6}$. При бросании двух игральных костей сумма выпадающего числа очков принимает свои возможные значения от 2 до 12 уже не одинаково часто; но относительные частоты этих возможных значений (при большом числе бросаний) будут близки к определенным числам, которые могут быть заранее подсчитаны по простым правилам (см. стр. 34). Установление подобных правил и решение других несколько более сложных задач, связанных с бросанием игральных костей, сыграли большую роль в начальный период развития теории вероятностей. И сейчас еще ряд основных понятий этой теории (случайное событие и его вероятность, случайная величина и др.) удобно иллюстрировать примерами с игральными костями. Так, бросание игральной кости и выпадение 1, 2, 3, 4, 5 или 6 очков на ее верхней грани наглядно поясняет понятие об испытании, имеющем ряд равновозможных исходов, и понятие о простейших случайных событиях; число очков, выпадающее на верхней грани игральной кости, служит простым примером случайной величины — величины, значения которой зависят от случая.

Но простейшие схемы, созданные для решения задач, связанных с бросанием игральных костей, имеют весьма ограниченную область применимости. Разнообразные задачи, поставленные перед теорией вероятностей развитием техники и естествознания в XIX и XX веках, потребовали изучения случайных величин значительно более сложной природы, особенно так называемых *непрерывных* случайных величин. Примерами таких величин являются, в частности, размер обрабатываемой детали и результат взвешивания, о которых мы говорили выше. Непрерывным случайным величинам и их числовым характеристикам будет уделено основное внимание в настоящем пособии.

В соответствии с указанной направленностью пособия первичные понятия случайного события и вероятности рассматриваются в нем весьма кратко (в первой главе).

Во второй главе подробно рассматриваются случайные величины двух наиболее важных типов (дискретные и непрерывные), а также функции от них. Изложение ограничивается в основном одномерными случайными величинами и ведется

отдельно для дискретных и отдельно для непрерывных случайных величин.

В третьей главе даются основные числовые характеристики случайных величин и доказываются их простейшие свойства. Важнейшие свойства средних значений, связанные с так называемым законом больших чисел, рассмотрены в главе четвертой. Пятая глава посвящена одному из центральных вопросов теории вероятностей — предельным теоремам, выясняющим роль так называемого нормального закона распределения случайных величин, в частности, для оценок средних значений.

В шестой главе рассматриваются некоторые приложения теории к математической обработке результатов измерений (или к теории ошибок измерения).

И, наконец, седьмая глава посвящена важному для практических приложений вопросу о линейной корреляции между случайными величинами.

ГЛАВА I

СЛУЧАЙНЫЕ СОБЫТИЯ И ВЕРОЯТНОСТИ

§ 1. Случайные события. Относительная частота и вероятность

Случайными событиями называются такие события, которые могут произойти или не произойти при осуществлении определенного комплекса условий *). Мы будем предполагать, что этот комплекс условий можно воспроизводить неограниченное количество раз. Каждое осуществление рассматриваемого комплекса условий назовем испытанием.

Пример: испытание — бросание игральной кости; случайное событие — выпадение шестерки, или выпадение четного числа очков. Другой пример: испытание — взвешивание некоторого тела на аналитических весах; случайное событие состоит в том, что ошибка измерения не превзойдет заранее заданного числа **).

Относительной частотой случайного события назовем отношение количества (m) случаев появления этого события к общему числу (n) произведенных испытаний. Опыт показывает, что *при многократном повторении испытания относительная частота* $\left(\frac{m}{n}\right)$ *случайного события обладает определенной устойчивостью:* если, например, при большом числе n испытаний относительная частота оказалась равной $\frac{m}{n}=0,2$, то и в любой другой серии из достаточно большого числа n' испытаний относительная частота $\frac{m'}{n'}$ будет близка к 0,2. Таким образом, в разных достаточно длинных

*) Разумеется, речь идет об условиях, существенным образом связанных с возможностью появления данных событий.

**) Ошибка измерения — разность между результатом взвешивания и истинным весом тела.

сериях испытаний относительные частоты случайного события как бы группируются около некоторого постоянного числа (разумеется, своего для каждого случайного события). Например, если игральная кость представляет собой точный куб, сделанный из однородного материала («правильная» игральная кость), то относительная частота выпадения 1, 2, 3, 4, 5 и 6 очков колеблется около одного и того же числа $\frac{1}{6}$.

Устойчивость относительной частоты может быть объяснена только как проявление некоторого объективного свойства случайного события, заключающегося в существовании определенной степени его возможности. Например, приблизительное равенство относительных частот выпадения 1, 2, 3, 4, 5 и 6 очков при бросании правильной игральной кости объясняется ее симметрией, делающей одинаково возможным выпадение каждого числа очков от 1 до 6.

Таким образом, *степень объективной возможности случайного события можно измерять числом. Это число называется вероятностью случайного события. Именно около этого числа группируются относительные частоты данного случайного события.* Как и относительная частота случайного события, его вероятность должна быть безразмерной величиной, заключенной между 0 и 1. Но в то время как относительная частота зависит еще и от произведенных испытаний, вероятность случайного события связана только с самим случайным событием (при данном комплексе условий).

Понятие вероятности является первичным, основным понятием, и в общем случае его нельзя определить через более простые понятия. Только в некоторых простейших схемах вероятность может быть подсчитана непосредственно, как будет показано в следующем параграфе; анализ таких простейших схем позволяет установить основные свойства вероятности, необходимые для дальнейшего построения курса.

§ 2. Классическое определение вероятности

Прежде всего, условимся об употреблении некоторых терминов. *Случайные события называются несовместимыми, если они не могут появиться одновременно. Случайные события образуют полную группу попарно несовместимых*

событий, если при каждом испытании должно появиться одно и только одно из них, то есть если каждые два из них несовместимы и хотя бы одно из них обязательно должно произойти.

В настоящем параграфе мы ограничимся рассмотрением *испытаний с равновозможными исходами;* примером может служить выпадение 1, 2, 3, 4, 5 и 6 очков на верхней грани правильной игральной кости *).

Точнее, мы будем здесь считать, что *исходы испытания можно представить в виде полной группы попарно несовместимых и равновозможных случайных событий*; эти случайные события будем коротко называть «случаями».

Если полная группа состоит из N равновозможных попарно несовместимых случаев, то каждому случаю приписывают вероятность, равную $\frac{1}{N}$. Это находится в согласии с тем, что при большом числе испытаний равновозможные случаи происходят приблизительно одинаково часто, то есть имеют относительную частоту, близкую к $\frac{1}{N}$. Например, при бросании правильной кости случаи выпадения 1, 2, 3, 4, 5 и 6 очков образуют полную группу; каждый случай будет иметь вероятность, равную $\frac{1}{6}$.

Рассмотрим теперь сложное событие A, заключающееся в появлении какого-либо из M фиксированных случаев. Вероятность случайного события A полагают по определению равной отношению $\frac{M}{N}$. Например, вероятность появления четного числа очков на верхней грани кости равна $\frac{3}{6}$, так как из шести указанных выше равновозможных случаев только в трех случаях появляется четное число очков (2, 4, 6).

Вероятность события A обозначают символом $\mathbf{P}\{A\}$.

*) При этом само понятие «равновозможности» мы уже не будем определять через более простые понятия. Оно основывается обычно на каких-либо соображениях симметрии, как в примере с игральной костью, и практически связано с равенством относительных частот всех исходов при большом числе испытаний. Отметим еще, что всюду в этом параграфе предполагается конечность числа рассматриваемых случаев.

Формула

$$\mathbf{P}\{A\} = \frac{M}{N} \qquad (1.1)$$

выражает так называемое классическое определение вероятности: *если результаты испытания можно представить в виде полной группы N равновозможных попарно несовместимых случаев и если случайное событие A появляется только в M случаях, то вероятность события A равна $\frac{M}{N}$, то есть равна отношению числа случаев, «благоприятствующих» событию, к общему числу всех случаев.*

Пример. При бросании двух монет герб может выпасть 2 раза, 1 раз или 0 раз (не выпасть ни разу); определить вероятности всех этих трех случайных событий (предполагается, что для каждой отдельной монеты выпадение или невыпадение герба равновозможны).

Перечисленные события образуют полную группу и, очевидно, попарно несовместимы. Но они не являются равновозможными. Для того чтобы применить классическое определение вероятности, надо представить все возможные исходы испытания в виде полной группы равновозможных событий. По соображениям симметрии это можно сделать следующим образом:

Первая монета	Вторая монета
Герб	Герб
Герб	Не герб
Не герб	Герб
Не герб	Не герб

Перечисленные здесь *четыре* исхода испытания естественно считать равновозможными, причем они снова составляют полную группу попарно несовместимых событий. Поэтому теперь уже можно применить классическое определение вероятности,

что дает:

— вероятность выпадения двух гербов равна $\frac{1}{4}$,

— вероятность выпадения одного герба равна $\frac{2}{4}=\frac{1}{2}$,

— вероятность невыпадения герба ни на одной монете равна $\frac{1}{4}$.

Подчеркнем еще раз, что классическое определение вероятности существенно опирается на предположение о равной возможности исходов испытания. Все задачи, к которым применимо классическое определение (1.1), укладываются в следующую простую схему — схему случайной выборки: из совокупности N элементов (предметов, явлений и т. п.) выбирается наудачу один элемент, причем каждому элементу обеспечивается одинаковая с остальными возможность быть выбранным; событие A заключается в выборе элемента, обладающего определенным признаком, причем этим признаком обладают точно M из N элементов рассматриваемой совокупности.

Проще всего представить себе осуществление этой схемы следующим образом: в урне находится N одинаковых на ощупь шаров, из них M шаров белых и $N-M$ не белых; испытание заключается в вынимании наугад одного шара из урны, случайное событие — в вынимании белого шара. При этих условиях вероятность вынуть белый шар равна $\frac{M}{N}$.

§ 3. Основные свойства вероятностей. Правило сложения вероятностей

Анализ данного в § 2 классического определения вероятности позволяет выявить следующие основные свойства вероятностей.

1. *Вероятность случайного события есть число неотрицательное:*

$$\mathbf{P}\{A\} \geqslant 0. \tag{1.2}$$

2. *Достоверное событие, то есть событие, которое при данном комплексе условий непременно должно произойти, имеет вероятность, равную единице:*

$$\mathbf{P}\{\text{достоверное событие}\} = 1. \tag{1.3}$$

3. Вероятности случайных событий подчиняются правилу сложения вероятностей: *если событие C состоит в осуществлении одного из двух несовместимых событий A или B (безразлично, какого именно), то вероятность события C равна сумме вероятностей событий A и B.*

Мы будем это правило записывать так:

$$\mathbf{P}\{A \text{ или } B\} = \mathbf{P}\{A\} + \mathbf{P}\{B\} \text{ (при несовместимости } A \text{ и } B). \quad (1.4)$$

Последнее свойство называют также *свойством аддитивности вероятностей.*

Первые два свойства непосредственно вытекают из формулы (1.1), так как в ней $M \geqslant 0$, $N > 0$ и для достоверного события $M = N$ (все исходы испытания благоприятствуют достоверному событию). Третье свойство доказывается для схемы случайной выборки следующим образом. Пусть в урне находится N шаров, из них K красных, L синих, остальные — белые; испытание заключается в вынимании из урны одного шара; событие A состоит в появлении красного шара, событие B — в появлении синего шара. Тогда событие (A или B) состоит в появлении цветного шара (безразлично, красного или синего). Непосредственный подсчет вероятностей по формуле (1.1) дает:

$$\mathbf{P}\{A\} = \frac{K}{N}; \qquad \mathbf{P}\{B\} = \frac{L}{N};$$

$$\mathbf{P}\{A \text{ или } B\} = \frac{K+L}{N},$$

что и доказывает формулу (1.4).

Для всех приложений теории вероятностей чрезвычайно важно то, что отмеченные свойства вероятностей справедливы не только для схемы случайной выборки, но и для любой системы случайных событий. Это утверждение можно обосновать следующим образом. Напомним, что общее понятие вероятности мы установили, исходя из факта устойчивости относительных частот случайных событий. Естественно поэтому считать, что основные свойства вероятностей случайных событий совпадают с основными свойствами относительных частот. Но для относительных частот указанные свойства легко проверить.

1*). Относительная частота $\frac{m}{n}$ не может быть отрицательной, так как $m \geqslant 0$, $n > 0$.

2*). Достоверное событие происходит при каждом повторении испытания, и поэтому его относительная частота равна $\frac{n}{n} = 1$.

3*). Если события A и B несовместимы, то событие (A или B) происходит столько раз, сколько раз происходит хотя бы одно из них (количество вынутых цветных шаров из описанной выше урны равно сумме количеств вынутых красных и синих шаров). Поэтому относительная частота события (A или B) равна сумме относительных частот событий A и B.

Исходя из приведенных выше соображений, *мы принимаем три указанных выше свойства вероятностей в качестве основных свойств вероятностей для любой системы случайных событий* *).

Замечание о предмете теории вероятностей. Теория вероятностей изучает не физическую сущность различных случайных событий, а лишь количественные соотношения между их вероятностями. Важную роль здесь играют основные свойства вероятностей и получаемые из них правила расчета. Дело в том, что типичной для теории вероятностей и ее приложений является следующая постановка задачи: имеется некоторая совокупность простых случайных событий, вероятности которых известны (заданы); требуется найти вероятности других случайных событий, связанных с данными событиями определенным образом. Например, при каждом бросании монеты вероятность выпадения герба принимается равной $\frac{1}{2}$; найти вероятность того, что при ста бросаниях монеты герб выпадет не менее 50 раз. Решение таких задач производится по определенным правилам расчета вероятностей (одно из таких правил — правило сложения — установлено выше). Для всех приложений совсем несущественно,

*) Полезно отметить, что на базе этих и некоторых других свойств можно аксиоматически построить всю теорию вероятностей; такое построение строго разработано известным советским математиком академиком А. Н. Колмогоровым (см. А. Н. Колмогоров, Основные понятия теории вероятностей, ОНТИ, 1936). Изложение этого вопроса выходит за рамки настоящего пособия.

как именно определяются вероятности исходной совокупности случайных событий; важно лишь, что если при достаточно большом числе испытаний относительные частоты исходных событий будут близки к их вероятностям, то это же будет верно и для относительной частоты интересующего нас сложного события, вероятность которого рассчитана по принятым нами правилам. Правила расчета, принятые в теории вероятностей, как раз отвечают этому основному требованию.

Следствия из основных свойств вероятностей

Следствие 1. Если случайные события $A_1, A_2, \ldots, A_n$ попарно несовместимы, то

$$\mathbf{P}\{A_1 \text{ или } A_2 \text{ или} \ldots \text{ или } A_n\} = \mathbf{P}\{A_1\} + \mathbf{P}\{A_2\} + \ldots + \mathbf{P}\{A_n\}. \quad (1.5)$$

Формула (1.5) может быть легко получена из формулы (1.4) методом полной индукции.

Следствие 2. *Если попарно несовместимые случайные события $A_1, A_2, \ldots, A_n$ образуют полную группу, то сумма их вероятностей равна единице.*

Действительно, для полной группы событий $A_1, A_2, \ldots, A_n$ событие (A_1 или A_2 или ... или A_n) является достоверным, и поэтому

$$\mathbf{P}\{A_1 \text{ или } A_2 \text{ или} \ldots \text{ или } A_n\} = 1.$$

Применяя к левой части этого равенства формулу (1.5), получаем:

$$\mathbf{P}\{A_1\} + \mathbf{P}\{A_2\} + \ldots + \mathbf{P}\{A_n\} = 1. \quad (1.6)$$

Особый интерес представляет частный случай, когда полная группа состоит только из двух несовместимых событий. При этом наступление одного из них равносильно ненаступлению другого. Такие случайные события называются взаимно противоположными. Если одно из пары взаимно противоположных событий обозначено через A, то другое обозначают через $\bar{A}$ (читают «не A»). *Вероятности двух взаимно противоположных событий в сумме дают единицу:*

$$\mathbf{P}\{A\} + \mathbf{P}\{\bar{A}\} = 1; \quad (1.7)$$

это непосредственно следует из формулы (1.6).

Таким образом, если известна вероятность какого-либо случайного события A, то вероятность противоположного ему события $\overline{A}$ вычисляется по формуле

$$\mathbf{P}\{\overline{A}\}=1-\mathbf{P}\{A\}.$$

Примером взаимно противоположных событий может служить выпадение герба и решетки при бросании монеты; если эти события равновозможны, то каждое из них имеет вероятность, равную $\frac{1}{2}$.

Невозможные события

Назовем невозможным событием такое событие, которое не может произойти ни при каком повторении испытания. Примером может служить вынимание белого шара из урны, где совсем нет белых шаров, или получение отрицательного результата при взвешивании тела. Невозможное событие можно считать противоположным любому достоверному событию (при том же комплексе условий); поэтому *вероятность невозможного события равна нулю.* Это согласуется с тем, что невозможное событие всегда имеет относительную частоту, равную нулю.

Заметим, что при классическом определении вероятность равна нулю тогда и только тогда, когда событие невозможно, то есть ни один возможный исход испытания не благоприятствует этому событию ($M=0$). В дальнейшем (при изучении непрерывных случайных величин) мы увидим, что из равенства нулю вероятности случайного события еще не следует его невозможность.

Связь классического определения с основными свойствами вероятностей

Как подчеркивалось выше, классическое определение вероятностей относится к такому случаю, когда исходы испытания можно представить в виде полной группы равновозможных событий. Интересно отметить, что в этом случае классическая формула (1.1) является единственной, которая согласуется с основными тремя свойствами вероятностей.

А именно, имеет место следующее предложение:

Если элементарные исходы испытания образуют полную группу равновозможных попарно несовместимых событий, то вероятность любого сложного события может быть выражена классической формулой (1.1).

Для доказательства обозначим элементарные исходы испытания через $E_1, E_2, \ldots, E_N$ и подсчитаем сначала их общую вероятность $p = \mathbf{P}\{E_k\}$ $(k = 1, 2, \ldots, N)$. По следствию 2 имеем:

$$\mathbf{P}\{E_1\} + \mathbf{P}\{E_2\} + \ldots + \mathbf{P}\{E_N\} = Np = 1,$$

откуда $p = \frac{1}{N}$. Обозначим, далее, через A сложное событие, которому благоприятствуют M фиксированных элементарных исходов испытания, например, $E_1, E_2, \ldots, E_M$. Тогда событие A есть сложное событие (E_1 или E_2 или ... или E_M) и по правилу сложения вероятностей

$$\mathbf{P}\{A\} = \mathbf{P}\{E_1\} + \mathbf{P}\{E_2\} + \ldots + \mathbf{P}\{E_M\},$$

то есть

$$\mathbf{P}\{A\} = Mp = \frac{M}{N}.$$

§ 4. Совмещение случайных событий. Независимые случайные события

Под совмещением случайных событий A и B понимают случайное событие, заключающееся в том, что в результате испытания произойдет и событие A и событие B. Совмещение случайных событий A и B мы будем обозначать через (A и B).

Пример. Из первой сотни чисел 1, 2, ..., 100 наудачу выбирается одно число; событие A заключается в том, что выбранное число делится на 3, событие B — в том, что оно делится на 4; тогда событие (A и B) заключается в том, что выбранное число делится и на 3 и на 4, то есть делится на 12. Легко подсчитать, что

$$\mathbf{P}\{A\} = \frac{33}{100}; \quad \mathbf{P}\{B\} = \frac{25}{100}; \quad \mathbf{P}\{A \text{ и } B\} = \frac{8}{100},$$

так как среди первых 100 чисел имеются 33 числа, делящихся на 3, 25 чисел, делящихся на 4, и 8 чисел, делящихся на 12.

Наиболее простое соотношение между вероятностями случайных событий A и B и вероятностью их совмещения (A и B) имеет место тогда, когда случайные события A и B ***независимы***. Понятие независимости мы поясним сначала на схеме случайной выборки. Пусть из двух урн с шарами вынимается наудачу по одному шару. Событие A заключается в том, что

шар, вынутый из первой урны, окажется белым, событие B — в том, что шар, вынутый из второй урны, окажется белым. Эти случайные события независимы по существу в том смысле, что цвет шара, вынутого из одной урны, не может влиять на цвет шара, вынутого из другой урны. Подсчитаем вероятность совмещения событий A и B, то есть вероятность того, что оба вынутых шара окажутся белыми. Обозначим количество шаров в первой и второй урнах соответственно через N_1 и N_2, а количество белых шаров в них соответственно через M_1 и M_2, так что

$$\mathbf{P}\{A\}=\frac{M_1}{N_1}; \quad \mathbf{P}\{B\}=\frac{M_2}{N_2}.$$

Так как каждый из N_1 исходов вынимания шара из первой урны может комбинироваться с каждым из N_2 исходов вынимания шара из второй урны, то число всех исходов равно N_1N_2; из них только в M_1M_2 случаях вынимаются два белых шара; следовательно, искомая вероятность совмещения равна

$$\mathbf{P}\{A \text{ и } B\}=\frac{M_1M_2}{N_1N_2},$$

то есть

$$\mathbf{P}\{A \text{ и } B\}=\mathbf{P}\{A\}\,\mathbf{P}\{B\}. \tag{1.8}$$

Формула (1.8) выражает правило умножения вероятностей для независимых случайных событий.

Напомним, что эта формула доказана нами только для частного случая. В общем случае само понятие независимости случайных событий нуждается в определении. Это может быть сделано с помощью формулы (1.8), простота которой делает ее важным инструментом в расчетах.

Определение. *Два случайных события A и B называются независимыми, если для них имеет место правило умножения вероятностей в форме* (1.8), *то есть если вероятность их совмещения равна произведению их вероятностей.*

Заметим, что из независимости случайных событий A и B следует попарная независимость событий $\bar{A}$ и B, A и $\bar{B}$, $\bar{A}$ и $\bar{B}$. Это утверждение может быть легко доказано формально, что мы предоставляем сделать читателю (см. упр. 5, стр. 28).

Данное выше определение независимости двух случайных событий можно расширить и на большее число событий:

события $A_1, A_2, \ldots, A_n$ называются независимыми в совокупности, если вероятность совмещения любых 2, 3, ... n из них равна произведению соответствующих вероятностей. Например, три события A, B, C независимы в совокупности, если имеют место четыре равенства:

$$\mathbf{P}\{A \text{ и } B\} = \mathbf{P}\{A\}\,\mathbf{P}\{B\};\ \mathbf{P}\{A \text{ и } C\} = \mathbf{P}\{A\}\,\mathbf{P}\{C\};$$

$$\mathbf{P}\{B \text{ и } C\} = \mathbf{P}\{B\}\,\mathbf{P}\{C\};$$

$$\mathbf{P}\{A \text{ и } B \text{ и } C\} = \mathbf{P}\{A\}\,\mathbf{P}\{B\}\,\mathbf{P}\{C\}. \qquad (1.9)$$

Полезно иметь в виду, что случайные события могут быть попарно независимы и все же не быть независимы в совокупности. Например, пусть испытание заключается в вынимании наудачу одного шара из урны, в которой лежат 4 шара с номерами 1, 2, 3 и 123, и пусть событие A (B, C) заключается в том, что на вынутом шаре окажется цифра 1 (соответственно, 2 и 3). События A, B и C попарно независимы, так как

$$\mathbf{P}\{A\} = \mathbf{P}\{B\} = \mathbf{P}\{C\} = \frac{2}{4} = \frac{1}{2};$$

$$\mathbf{P}\{A \text{ и } B\} = \mathbf{P}\{B \text{ и } C\} = \mathbf{P}\{C \text{ и } A\} = \frac{1}{4} = \frac{1}{2}\cdot\frac{1}{2}.$$

Но события A, B и C не будут независимы в совокупности, так как

$$\mathbf{P}\{A \text{ и } B \text{ и } C\} = \frac{1}{4} \neq \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}.$$

Отметим также, что наличие одного только соотношения (1.9) еще не обеспечивает независимости случайных событий A, B, C в совокупности. Например, если в урне лежит 8 шаров с номерами 1, 2, 3, 12, 13, 20, 30, 123 и если события A, B, C имеют тот же смысл, что и в предыдущем примере, то

$$\mathbf{P}\{A\} = \mathbf{P}\{B\} = \mathbf{P}\{C\} = \frac{4}{8} = \frac{1}{2};\ \mathbf{P}\{A \text{ и } B \text{ и } C\} = \frac{1}{8} = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2};$$

и даже

$$\mathbf{P}\{A \text{ и } B\} = \frac{2}{8} = \frac{1}{2}\cdot\frac{1}{2},\ \mathbf{P}\{A \text{ и } C\} = \frac{2}{8} = \frac{1}{2}\cdot\frac{1}{2};$$

но

$$\mathbf{P}\{B \text{ и } C\} = \frac{1}{8} \neq \mathbf{P}\{B\}\cdot\mathbf{P}\{C\}.$$

Обобщение правила сложения вероятностей

Если случайные события A и B независимы, то

$$\mathbf{P}\{A \text{ или } B\} = \mathbf{P}\{A\} + \mathbf{P}\{B\} - \mathbf{P}\{A\}\mathbf{P}\{B\}. \qquad (1.10)$$

Для доказательства формулы (1.10) заметим, прежде всего, что события (A или B) и ($\bar{A}$ и $\bar{B}$) взаимно противоположны (если происходит хотя бы одно из двух событий A или B, то не происходит соответствующее противоположное событие, значит, не может произойти и совмещение противоположных событий $\bar{A}$ и $\bar{B}$). Применяя формулу (1.7) и правило умножения, получаем:

$$\mathbf{P}\{A \text{ или } B\} = 1 - \mathbf{P}\{\bar{A} \text{ и } \bar{B}\} = 1 - \mathbf{P}\{\bar{A}\}\mathbf{P}\{\bar{B}\} =$$
$$= 1 - (1 - \mathbf{P}\{A\})(1 - \mathbf{P}\{B\}) = \mathbf{P}\{A\} + \mathbf{P}\{B\} - \mathbf{P}\{A\}\mathbf{P}\{B\},$$

что и требовалось доказать.

Приведем еще без доказательства общее правило сложения вероятностей:

$$\mathbf{P}\{A \text{ или } B\} = \mathbf{P}\{A\} + \mathbf{P}\{B\} - \mathbf{P}\{A \text{ и } B\}. \qquad (1.11)$$

(Здесь события A и B не обязательно независимы.)

Упражнения. 1. Доказать формулу (1.11) в случае классического определения вероятностей.

2. Дать геометрическое толкование формулы (1.11), считая испытанием бросание точки на единичный квадрат и принимая при этом вероятность попадания точки в некоторую область внутри квадрата равной площади этой области (рис. 1).

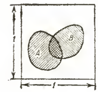


Рис. 1.

Пример 1. Два стрелка независимо друг от друга стреляют по одной и той же цели; вероятность попадания для первого стрелка равна $\mathbf{P}\{A\} = 0,9$, для второго $\mathbf{P}\{B\} = 0,8$.

Требуется определить вероятность поражения цели, то есть вероятность того, что хотя бы один стрелок попадет в цель. По формуле (1.10) находим

$$P\{A \text{ или } B\} = 0,9 + 0,8 - 0,9 \cdot 0,8 = 0,98.$$

Пример 2. Некоторое количество n стрелков независимо друг от друга стреляют по одной и той же цели; вероят-

ность попадания для каждого стрелка равна p. Определить количество стрелков, которое требуется для поражения цели с вероятностью не меньшей, чем P.

Для этого найдем зависимость между p и P. Вероятность непоражения цели одним стрелком равна $1-p$; вероятность того, что ни один стрелок не попадет в цель, равна $(1-p)^n$; события «ни один стрелок не попадет в цель» и «хотя бы один стрелок попадет в цель» (то есть «цель будет поражена») взаимно противоположны, поэтому вероятность поражения цели есть $1-(1-p)^n$. По условию

$$1-(1-p)^n \geqslant P.$$

Отсюда находим n — целое число, удовлетворяющее неравенству

$$n \geqslant \frac{\lg(1-P)}{\lg(1-p)}.$$

Например, если при стрельбе по самолету вероятность попадания равна $p=0{,}004$, то для обеспечения поражения самолета с вероятностью не меньшей $P=0{,}98$ требуется

$$n \geqslant \frac{\lg 0{,}02}{\lg 0{,}996}, \text{ или } n \geqslant 976 \text{ стрелков.}$$

§ 5. Условные вероятности. Общее правило умножения вероятностей. Формула полной вероятности

Целью настоящего параграфа является обобщение правила умножения вероятностей (1.8) на зависимые случайные события. Для этого обратимся сначала, как и в § 4, к классической схеме случайной выборки. Пусть испытание заключается в вынимании наугад одного шара из урны, содержащей N шаров, одинаковых на ощупь, но различающихся по двум признакам: по цвету и по рисунку. Введем обозначения:

K — количество цветных шаров ($N-K$ — белых),

L — количество шаров с рисунком ($N-L$ — без рисунка),

M — количество цветных шаров с рисунком.

Пусть событие A заключается в появлении цветного шара, событие B — в появлении шара с рисунком. Совмещение событий A и B означает появление цветного шара с рисунком.

Вероятности этих случайных событий соответственно равны

$$\mathbf{P}\{A\}=\frac{K}{N}; \quad \mathbf{P}\{B\}=\frac{L}{N}; \quad \mathbf{P}\{A \text{ и } B\}=\frac{M}{N}.$$

Попробуем, по аналогии с формулой (1.8), связать вероятность события (A и B) с вероятностью события A; мы получим:

$$\frac{M}{N}=\frac{K}{N}\,\frac{M}{K}. \tag{1.12}$$

Отношение $\frac{M}{K}$ количества цветных шаров с рисунком к количеству всех цветных шаров также имеет характер вероятности, а именно, оно дает вероятность выбрать шар с рисунком при условии, что выбор производится только из числа цветных шаров (ибо из всех K цветных шаров только M шаров имеют рисунок). Такую вероятность называют *условной вероятностью события B при условии осуществления события A*; обозначают эту условную вероятность через $\mathbf{P}\{B|A\}$, так что в нашем примере

$$\mathbf{P}\{B|A\}=\frac{M}{K}.$$

Теперь мы можем записать соотношение (1.12) в виде

$$\mathbf{P}\{A \text{ и } B\}=\mathbf{P}\{A\}\,\mathbf{P}\{B|A\}. \tag{1.13}$$

Это соотношение выражает общее правило умножения вероятностей: *вероятность совмещения двух случайных событий равна произведению вероятности одного из них на условную вероятность другого.*

Формула (1.13) выведена нами для классической схемы.

Обратимся теперь к общему случаю каких угодно случайных событий A и B. Здесь формула (1.13) служит для определения условной вероятности. А именно, *условная вероятность события B при условии осуществления события A определяется с помощью формулы*

$$\mathbf{P}\{B|A\}=\frac{\mathbf{P}\{A \text{ и } B\}}{\mathbf{P}\{A\}} \quad (\text{при } \mathbf{P}\{A\}\neq 0). \tag{1.14}$$

Точно так же вводится условная вероятность события A при условии осуществления события B:

$$\mathbf{P}\{A|B\}=\frac{\mathbf{P}\{A \text{ и } B\}}{\mathbf{P}\{B\}} \quad (\text{при } \mathbf{P}\{B\}\neq 0). \tag{1.15}$$

Условные вероятности, как легко проверить, обладают всеми основными свойствами вероятностей.

Формула (1.13) может быть обобщена и на большее число случайных событий. Например, для трех случайных событий A, B, C

$$\mathbf{P}\{A \text{ и } B \text{ и } C\} = \mathbf{P}\{A \text{ и } B\}\,\mathbf{P}\{C \mid A \text{ и } B\} = \mathbf{P}\{A\}\,\mathbf{P}\{B \mid A\}\,\mathbf{P}\{C \mid A \text{ и } B\}.$$

Введение понятия условной вероятности позволяет дать новое толкование независимости случайных событий. Если случайные события A и B независимы, то из формул (1.8), (1.14) и (1.15) следует:

$$\mathbf{P}\{B \mid A\} = \frac{\mathbf{P}\{A\}\,\mathbf{P}\{B\}}{\mathbf{P}\{A\}} = \mathbf{P}\{B\},$$

$$\mathbf{P}\{A \mid B\} = \frac{\mathbf{P}\{A\}\,\mathbf{P}\{B\}}{\mathbf{P}\{B\}} = \mathbf{P}\{A\},$$

то есть следует равенство условных и безусловных вероятностей этих событий.

Очевидно также, что и обратно, при $\mathbf{P}\{B \mid A\} = \mathbf{P}\{B\}$ формула (1.13) превращается в формулу (1.8).

Таким образом, независимость случайных событий A и B означает, что вероятность события B (или A) не изменяется при введении дополнительного условия осуществления события A (или соответственно B) *).

Из этого толкования непосредственно следует, например, что достоверное событие и любое случайное событие A независимы.

В рассмотренном выше примере независимость случайных событий A и B сводится к выполнению равенства

$$\frac{M}{K} = \frac{L}{N},$$

то есть к условию, что доля шаров с рисунком среди цветных шаров равна доле шаров с рисунком среди всех шаров

*) Напомним, что вероятность случайного события (B) всегда связана с определенным комплексом условий. Если к этому комплексу добавить дополнительное условие осуществления некоторого другого события (A), то вероятность рассматриваемого события (B) может измениться.

в урне. Такое пропорциональное распределение шаров по двум признакам представлено схематически на рис. 2, *а*, где $\mathbf{P}\{A\}=0{,}4$; $\mathbf{P}\{B\}=\mathbf{P}\{B|A\}=0{,}3$. Для сравнения на

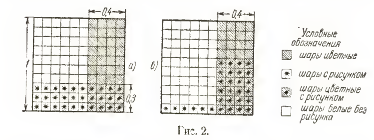


Рис. 2.

рис. 2, *б* дан пример непропорционального распределения шаров по двум признакам, где

$$\mathbf{P}\{A\}=0{,}4;\quad \mathbf{P}\{B\}=\frac{24+6}{100}=0{,}3;\quad \text{но } \mathbf{P}\{B|A\}=\frac{24}{40}=0{,}6.$$

Формула полной вероятности

Теорема. *Если случайные события $H_1, H_2, \ldots, H_n$ попарно несовместимы и если событие A может осуществиться только с каким-нибудь одним из этих событий, то*

$$\mathbf{P}\{A\}=\mathbf{P}\{H_1\}\mathbf{P}\{A|H_1\}+\mathbf{P}\{H_2\}\mathbf{P}\{A|H_2\}+\ldots \\ \ldots+\mathbf{P}\{H_n\}\mathbf{P}\{A|H_n\}. \quad (1.16)$$

Доказательство. При высказанных условиях событие A равносильно совмещению событий

$$(H_1 \text{ или } H_2 \text{ или } \ldots \text{ или } H_n) \text{ и } A.$$

Но это совмещение происходит тогда и только тогда, когда происходит одно из совмещений

$$(H_1 \text{ и } A), \text{ или } (H_2 \text{ и } A), \text{ или } \ldots, \text{ или } (H_n \text{ и } A).$$

Применяя правило сложения вероятностей, получаем:

$$\mathbf{P}\{A\}=\mathbf{P}\{(H_1, \text{ или } H_2, \text{ или } \ldots \text{ или } H_n) \text{ и } A\}= \\ =\mathbf{P}\{H_1 \text{ и } A\}+\mathbf{P}\{H_2 \text{ и } A\}+\ldots+\mathbf{P}\{H_n \text{ и } A\}; \quad (1.17)$$

остается лишь применить общее правило умножения вероятностей:

$$\mathbf{P}\{H_1 \text{ и } A\}=\mathbf{P}\{H_1\}\mathbf{P}\{A|H_1\};\ \ldots$$

В частности, всегда имеет место формула

$$\mathbf{P}\{B\} = \mathbf{P}\{A\}\,\mathbf{P}\{B\mid A\} + \mathbf{P}\{\bar{A}\}\,\mathbf{P}\{B\mid\bar{A}\}, \qquad (1.18)$$

так как противоположные события A и $\bar{A}$ несовместимы и образуют полную группу.

Пример 1. В урне находится N шаров, из них M белых; испытание заключается в том, что из урны последовательно вынимаются два шара; событие A заключается в том, что первый вынутый шар окажется белым, событие B — в том, что второй вынутый шар окажется белым.

Очевидно,

$$\mathbf{P}\{A\} = \frac{M}{N}; \quad \mathbf{P}\{\bar{A}\} = \frac{N-M}{N};$$

$$\mathbf{P}\{B\mid A\} = \frac{M-1}{N-1}; \quad \mathbf{P}\{B\mid\bar{A}\} = \frac{M}{N-1}.$$

Подсчитаем $\mathbf{P}\{B\}$. По формуле (1.18):

$$\mathbf{P}\{B\} = \frac{M}{N}\,\frac{M-1}{N-1} + \frac{N-M}{N}\,\frac{M}{N-1} = \frac{M}{N}.$$

Таким образом, вероятность того, что второй вынутый шар будет белым, равна вероятности того, что будет белым первый вынутый шар.

Пример 2. В доске имеются отверстия (ячейки) с координатами $(x_k;\ y_l)$ $(k = 1, 2, \ldots, n;\ l = 1, 2, \ldots, m)$. На доску брошен шарик, который может попасть в одну из ячеек. Вероятности попадания шарика в каждую из ячеек приведены в таблице:

$y \backslash x$	x_1	x_2	...	x_n
y_1	p_{11}	p_{21}	...	p_{n1}
y_2	p_{12}	p_{22}	...	p_{n2}
...	...	...	...	...
y_m	p_{1m}	p_{2m}	...	p_{nm}

(1.19)

Здесь p_{kl} есть вероятность попадания шарика в ячейку с координатами $(x_k;\ y_l)$. Вычислим вероятность P_k попадания шарика в ячейку с абсциссой x_k.

Так как ячейка с абсциссой x_k может иметь одну и только одну из ординат $y_1, y_2, \ldots, y_m$, то по формуле (1.17) получаем:

$$P_k = p_{k1} + p_{k2} + \ldots + p_{km} \qquad (k = 1,\ 2,\ \ldots,\ n). \quad (1.20)$$

Упражнения

1. Из колоды карт (52 карты) берется 3 карты. Вычислить вероятность того, что среди взятых карт будет хотя бы один туз.

Ответ. $p = \frac{1201}{5525} = 0{,}217$.

2. Рабочий обслуживает три станка, работающих независимо друг от друга. Вероятность того, что в течение часа станок не потребует внимания рабочего, равна для первого станка 0,9, для второго 0,8 и для третьего 0,7. Вычислить вероятность того, что по крайней мере один из трех станков не потребует внимания рабочего в течение часа.

Ответ. $p = 0{,}994$.

3. В урне находится 5 белых и 20 черных шаров. Из урны последовательно вынимаются шары до тех пор, пока не будет вынут белый шар. Вычислить вероятность того, что при этих условиях будет произведено три вынимания, то есть что до первого белого шара будет вынуто 2 черных шара.

Ответ. $p = \frac{19}{138} = 0{,}138$.

4. На двух станках обрабатываются однотипные детали; вероятность брака для станка № 1 составляет 0,03, а для станка № 2 — 0,02. Обработанные детали складываются в одном месте, причем деталей со станка № 1 складывается вдвое больше, чем со станка № 2. Вычислить вероятность того, что взятая наудачу деталь не будет бракованной.

Ответ. $p = \frac{292}{300} = 0{,}973$.

5. С помощью формулы (1.18) доказать, что из независимости случайных событий A и B следует независимость случайных событий $\bar{A}$ и B.

ГЛАВА II

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

§ 6. Дискретные случайные величины

В настоящем параграфе предметом изучения являются величины, значения которых зависят от случая, как, например, число очков, выпадающее на верхней грани игральной кости, или число вызовов на телефонной станции за данный промежуток времени. Те значения, которые в результате испытания может принять изучаемая величина, будем называть ее возможными значениями.

Определение. *Величина* ξ *называется дискретной случайной величиной, если все ее возможные значения образуют конечную или бесконечную последовательность чисел* $x_1, x_2, \ldots, x_k, \ldots$ *и если принятие ею каждого из указанных значений есть случайное событие с определенной вероятностью.*

Вероятность случайного события $(\xi = x_k)$ обозначим через p_k и будем говорить, что p_k есть вероятность значения x_k. Вероятность p_k есть функция от x_k. Эта функция называется *законом распределения вероятностей величины* ξ.

Всякое правило, позволяющее находить вероятности всех возможных значений величины ξ, *определяет закон распределения ее вероятностей*. Обычно этот закон записывают в виде таблицы, в которой перечисляются все возможные (различные) значения величины ξ и их вероятности:

Возможное значение ξ	x_1	x_2	...	x_k	...
Вероятность (p)	p_1	p_2	...	p_k	...

Такая таблица называется *таблицей распределения вероятностей* дискретной случайной величины ξ.

Если случайная величина ξ может принимать лишь конечное число различных значений $x_1, x_2, \ldots, x_n$, то случайные события

$$(\xi = x_1),\ (\xi = x_2),\ \ldots,\ (\xi = x_n)$$

образуют полную группу попарно несовместимых событий; следовательно, сумма их вероятностей должна быть равна 1:

$$p_1 + p_2 + \ldots + p_n = 1. \qquad (2.1)$$

Если таблица распределения вероятностей содержит бесконечно много значений, то условие (2.1) заменяется следующим: бесконечный ряд $p_1 + p_2 + \ldots + p_k + \ldots$ должен быть сходящимся и его сумма должна быть равна 1.

Пример 1. Число очков, выпадающее на верхней грани правильной игральной кости, есть дискретная случайная величина со следующей таблицей распределения вероятностей (см. § 2):

Число очков	1	2	3	4	5	6
Вероятность	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

(2.2)

Заметим, что для неправильной игральной кости возможные значения числа очков остаются теми же, но вероятности их могут быть отличны от $\frac{1}{6}$.

Пример 2. Охотник, имеющий три патрона, стреляет в цель до первого попадания (или пока не израсходует все три патрона). Число израсходованных патронов будет случайной величиной (ξ) с тремя возможными значениями (1, 2, 3). Найдем распределение вероятностей этой величины при условии, что вероятность попадания при каждом выстреле равна 0,8.

Событие ($\xi = 1$) означает попадание с первого выстрела, поэтому его вероятность равна

$$\mathbf{P}\{\xi = 1\} = 0{,}8.$$

Событие ($\xi = 2$) означает попадание лишь со второго выстрела

(и, значит, непопадание при первом выстреле), его вероятность равна

$$\mathbf{P}\{\xi=2\}=(1-0{,}8)\cdot 0{,}8=0{,}2\cdot 0{,}8=0{,}16.$$

Наконец, три выстрела производятся, если не было попадания при первых двух выстрелах, поэтому

$$\mathbf{P}\{\xi=3\}=0{,}2\cdot 0{,}2=0{,}04.$$

Эту же вероятность можно было бы подсчитать и по формуле (2.1):

$$\mathbf{P}\{\xi=3\}=1-\mathbf{P}\{\xi=1\}-\mathbf{P}\{\xi=2\}=$$
$$=1-0{,}8-0{,}16=0{,}04.$$

Итак, таблица распределения вероятностей величины ξ есть

ξ	1	2	3
p	0,8	0,16	0,04

(2.3)

Пример 3. Производится стрельба по некоторой цели до первого попадания, без ограничения числа выстрелов. Вероятность попадания при каждом выстреле равна p. Число произведенных выстрелов есть случайная величина ξ с бесконечной таблицей распределения вероятностей:

ξ	1	2	3	...	n	...
p	$(1-p)\,p$	$(1-p)^2\,p$	...	$(1-p)^{n-1}p$	...

(2.4)

Ряд вероятностей представляет собой бесконечно убывающую геометрическую прогрессию со знаменателем $(1-p)$: он сходится, и его сумма равна

$$p+(1-p)\,p+(1-p)^2p+\ldots+(1-p)^{n-1}p+\ldots=$$
$$=\frac{p}{1-(1-p)}=1.$$

Пример 4. В некоторых задачах физики и техники встречаются случайные величины, подчиненные закону распределения Пуассона

ξ	0	1	2	$\ldots$	m	$\ldots$
	e^{-a}	ae^{-a}	$\frac{a^2}{2!}e^{-a}$	$\ldots$	$\frac{a^m}{m!}e^{-a}$	$\ldots$

, (2.5)

где a — некоторое положительное число, характеризующее случайную величину ξ.

Закону распределения Пуассона подчиняются, например: а) количество вызовов на автоматической телефонной станции за данный промежуток времени; б) количество электронов, вылетающих с накаленного катода за данный промежуток времени. Что представляет собой число a в этих примерах, будет ясно из дальнейшего. Здесь же мы отметим только, что ряд вероятностей $\sum\limits_{m=0}^{\infty}\frac{a^m}{m!}e^{-a}$ сходится и его сумма равна 1:

$$e^{-a}\left(1+a+\frac{a^2}{2!}+\ldots+\frac{a^m}{m!}+\ldots\right)=e^{-a}e^{+a}=1.$$

Линейные действия над случайными величинами

К линейным действиям мы будем относить умножение случайной величины ξ на число и сложение случайных величин.

Произведением $C\xi$ *дискретной случайной величины*

ξ	x_1	x_2	$\ldots$
	p_1	p_2	$\ldots$

на число C называется дискретная случайная величина с распределением вероятностей

$C\xi$	Cx_1	Cx_2	$\ldots$
	p_1	p_2	$\ldots$

(2.6)

Иначе говоря, умножение дискретной случайной величины на число сводится только к умножению всех ее значений на это число (без изменения вероятностей).

Например, если охотник в примере 2 (стр. 30) платит 2 рубля за каждый израсходованный патрон, то истраченная им сумма (в рублях) будет случайной величиной со следующим распределением вероятностей:

2ξ	2	4	6
p	0,8	0,16	0,04

Несколько сложнее определяется распределение вероятностей суммы $\xi+\eta$ двух дискретных величин

ξ	x_1	x_2	...
	p_1	p_2	...

,

η	y_1	y_2	...
	q_1	q_2	...

Обозначим через p_{kl} вероятность совмещения случайных событий $(\xi=x_k)$ и $(\eta=y_l)$. Если происходит указанное совмещение, то сумма $\xi+\eta$ принимает значение x_k+y_l. Однако вероятность, например, события $(\xi+\eta=x_1+y_1)$ может оказаться больше, чем p_{11}, если среди сумм x_k+y_l встретятся числа, равные x_1+y_1. А именно, по правилу сложения вероятностей мы должны считать вероятность события $(\xi+\eta=x_1+y_1)$ равной сумме всех тех вероятностей p_{kl}, для которых число x_k+y_l равно x_1+y_1. Таким образом, *значениями суммы $\xi+\eta$ служат суммы всех возможных значений величин ξ и η, а вероятность каждого из указанных значений суммы равна сумме вероятностей тех совмещений $(\xi=x_k)$ и $(\eta=y_l)$, при которых достигается данное значение суммы.*

На практике для построения таблицы распределения суммы $\xi+\eta$ обычно строят сначала вспомогательную таблицу

x_1+y_1	x_1+y_2	x_2+y_1	x_2+y_2	...
p_{11}	p_{12}	p_{21}	p_{22}	...

, (2.7)

а затем объединяют в ней равные значения x_k+y_l, складывая соответствующие вероятности p_{kl}. Заметим, что если все

значения $x_k + y_l$ различны, то таблица (2.7) будет готовой таблицей распределения суммы $\xi + \eta$.

Пример. Пусть испытание заключается в одновременном бросании двух правильных игральных костей;

ξ — число очков, выпадающее на первой кости;

η — число очков, выпадающее на второй кости;

$\xi + \eta$ — сумма чисел очков, выпадающих на двух костях.

Обе случайные величины ξ и η имеют одинаковые таблицы распределения вероятностей (2.2). Найдем распределение вероятностей их суммы $\xi + \eta$. Так как выпадение любого числа очков на первой кости не зависит от выпадения любого числа очков на второй кости, то вероятность каждого совмещения будет равна $\frac{1}{36}$. По правилу (2.7) построим сначала вспомогательную таблицу

$1+1$	$1+2$	$2+1$	$1+3$	$2+2$	$3+1$	...	$6+6$
$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	...	$\frac{1}{36}$

Объединяя равные числа в первой строке — строке возможных значений, получаем таблицу распределения вероятностей величины $(\xi + \eta)$:

2	3	4	5	6	7	8	9	10	11	12
$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Отметим следующую особенность сложения случайных величин; сравнивая таблицу распределения для суммы $\xi + \eta$ с таблицей распределения для 2ξ:

2	4	6	8	10	12
$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

мы приходим к выводу, что сложение случайных величин с одинаковыми распределениями вероятностей не сводится, во-

обще говоря, к умножению одной из них на целое число. Введенные выше действия сложения случайных величин и умножения их на число сохраняют известные свойства сложения и умножения чисел. В частности, имеют место следующие легко проверяемые формулы:

$$\xi+\eta=\eta+\xi; \qquad (\xi+\eta)+\zeta=\xi+(\eta+\zeta);$$
$$C(\xi+\eta)=C\xi+C\eta.$$

Независимость случайных величин

Дискретные случайные величины ξ, η называются *независимыми*, если независимы случайные события $(\xi=x_k)$ и $(\eta=y_l)$ при всех k и l, то есть если вероятности совмещения этих событий находятся по правилу умножения вероятностей:

$$p_{kl}=p_k q_l \qquad (k=1,\ 2,\ \ldots;\ l=1,\ 2,\ \ldots). \quad (2.8)$$

Например, при бросании двух игральных костей числа очков, выпадающих на первой и второй костях, являются независимыми случайными величинами; это упростило нахождение их суммы в рассмотренном выше примере.

Случайные величины ξ_1, ξ_2, ..., ξ_n называются *взаимно независимыми*, если независимы в совокупности все случайные события $(\xi_1=x_{k_1}^{(1)})$, $(\xi_2=x_{k_2}^{(2)})$, ..., $(\xi_n=x_{k_n}^{(n)})$ (здесь через $x_1^{(i)}$, $x_2^{(i)}$, ..., $x_{k_i}^{(i)}$, ... обозначены значения случайной величины ξ_i).

Если случайные величины ξ_1, ξ_2, ..., ξ_n взаимно независимы, то по распределениям их вероятностей легко найти распределение вероятностей любой их линейной комбинации $C_1\xi_1+C_2\xi_2+\ldots+C_n\xi_n$ с постоянными коэффициентами C_1, C_2, ..., C_n.

Этим обстоятельством часто пользуются в расчетах, представляя изучаемую случайную величину в виде линейной комбинации независимых случайных величин с известными распределениями вероятностей.

§ 7. Распределение вероятностей относительной частоты случайного события

Рассмотрим относительную частоту ω_n случайного события A при n-кратном повторении испытания. Будем считать, что появление события A в каждом испытании не зависит от его появления в других испытаниях и что вероятность

события A в каждом испытании одна и та же. Обозначим ее через p *).

Так как за n испытаний событие A может произойти 0, 1, 2, ..., n раз, то ω_n будет дискретной случайной величиной с возможными значениями $0, \frac{1}{n}, \frac{2}{n}, \ldots, 1$. Найдем распределение ее вероятностей. С этой целью представим величину ω_n в виде линейной комбинации более простых случайных величин. Введем так называемую *характеристическую случайную величину* λ_k — число появлений события A при k-м испытании. Величина λ_k может принимать только два значения: 1, если событие A произойдет при k-м испытании, и 0, если событие A не произойдет при k-м испытании. Так как вероятность события A равна p в каждом испытании, то величины $\lambda_1, \lambda_2, \ldots, \lambda_n$ имеют одинаковые таблицы распределения вероятностей:

$$\lambda_1 \begin{array}{|c|c|}\hline 1 & 0 \\ \hline p & q \\ \hline\end{array}, \quad \lambda_2 \begin{array}{|c|c|}\hline 1 & 0 \\ \hline p & q \\ \hline\end{array}, \ldots, \lambda_n \begin{array}{|c|c|}\hline 1 & 0 \\ \hline p & q \\ \hline\end{array}, \qquad (2.9)$$

где $q = 1 - p$. Все величины $\lambda_1, \lambda_2, \ldots, \lambda_n$ взаимно независимы в силу принятого нами условия.

Рассмотрим теперь сумму характеристических величин

$$\mu_n = \lambda_1 + \lambda_2 + \ldots + \lambda_n. \qquad (2.10)$$

Эта сумма состоит из единиц и нулей, причем единиц в ней

*) Схематически это повторение испытаний можно представить следующим образом. В урну помещается определенное количество одинаковых на ощупь шаров, часть из которых помечена меткой «A» (например, белым цветом); доля меченых шаров должна быть равна p, так что вероятность вынуть меченый шар равна вероятности события A. Вынув из урны наугад один шар, мы записываем, есть ли на нем метка или нет, возвращаем шар в урну, тщательно перемешиваем шары и затем снова вынимаем один шар. Этот процесс повторяется до получения n записей. Такая последовательность испытаний называется *последовательностью независимых испытаний по схеме Бернулли или по схеме возвращенного шара*. Ее называют также *схемой повторной выборки* (в отличие от бесповторной выборки, при которой вынутый шар не возвращают в урну).

ровно столько, сколько раз произойдет событие A за n испытаний; следовательно, величина μ_n равна количеству повторений события A за n испытаний, а отношение μ_n к общему числу испытаний (n) равно относительной частоте ω_n:

$$\omega_n = \frac{\mu_n}{n} = \frac{1}{n}\left(\lambda_1 + \lambda_2 + \ldots + \lambda_n\right). \tag{2.11}$$

Полученное представление случайной величины ω_n в виде линейной комбинации взаимно независимых случайных величин $\lambda_1, \lambda_2, \ldots, \lambda_n$ с известными распределениями вероятностей (2.9) позволяет найти распределение вероятностей для ω_n с помощью формул предыдущего параграфа.

Будем складывать величины $\lambda_1, \lambda_2, \ldots, \lambda_n$ последовательно. Прежде всего, по формулам (2.7) и (2.8) имеем:

λ_1

1	0
p	q

$+\lambda_2$

1	0
p	q

$=\mu_2$

2	1	0
pp	$pq+qp$	qq

,

то есть

μ_2

2	1	0
p^2	$2pq$	q^2

Далее находим тем же способом

$\mu_2 + \lambda_3 = \mu_3$

3	2	1	0
p^2p	$p^2q+2pqp$	$2pqq+q^2p$	q^2q

,

то есть

μ_3

3	2	1	0
p^3	$3p^2q$	$3pq^2$	q^3

Мы замечаем, что вероятности в таблицах распределения величин μ_2 и μ_3 совпадают с соответствующими членами разложения биномов

$$(p+q)^2 = p^2 + 2pq + q^2;$$
$$(p+q)^3 = p^3 + 3p^2q + 3pq^2 + q^3$$

(откуда, кстати, сразу видно, что суммы вероятностей в этих таблицах равны единице, так как $p+q=1$).

Методом математической индукции можно доказать следующее общее утверждение: *вероятность того, что* μ_n *примет некоторое значение* m, *равна члену, содержащему* p^m, *в разложении бинома* $(p+q)^n$ *по степеням* p:

$$\mathbf{P}\{\mu_n = m\} = C_n^m p^m q^{n-m} = \frac{n!}{m!\,(n-m)!} p^m q^{n-m}. \quad (2.12)$$

Справедливость формулы (2.12) при $n=2$ (а также и при $n=3$) была уже установлена раньше. Докажем, что из справедливости этой формулы при некотором числе n испытаний вытекает ее справедливость и при $n+1$ испытаниях. Действительно, величина $\mu_{n+1} = \mu_n + \lambda_{n+1}$ может принять значение m только в двух случаях: либо при $\mu_n = m$ и $\lambda_{n+1} = 0$, либо при $\mu_n = m-1$ и $\lambda_{n+1} = 1$; поэтому по правилам сложения и умножения вероятностей

$$\mathbf{P}\{\mu_{n+1} = m\} = \mathbf{P}\{\mu_n = m\}\,\mathbf{P}\{\lambda_{n+1} = 0\} + \mathbf{P}\{\mu_n = m-1\}\,\mathbf{P}\{\lambda_{n+1} = 1\}.$$

По формулам (2.12) и (2.9) получаем:

$$\mathbf{P}\{\mu_{n+1} = m\} = \frac{n!}{m!\,(n-m)!} p^m q^{n-m} q + \frac{n!}{(m-1)!\,(n-m+1)!} p^{m-1} q^{n-m+1} p,$$

то есть

$$\mathbf{P}\{\mu_{n+1} = m\} = \frac{(n+1)!}{m!\,(n+1-m)!} p^m q^{n+1-m},$$

что и требовалось доказать.

Формулу (2.12) можно объяснить и не прибегая к представлению величины μ_n в виде (2.10). Вероятность того, что событие A наступит в первых m испытаниях и не наступит в остальных $n-m$ испытаниях, может быть подсчитана по правилу умножения вероятностей для независимых событий, что дает:

$$p^m q^{n-m}. \quad (2.13)$$

Эта вероятность не зависит от того, в каких именно m испытаниях наступит событие A. Но из n последовательных

испытаний можно C_n^m различными способами выбрать m испытаний, в которых будет иметь место событие A. Поэтому, по правилу сложения вероятностей, искомая вероятность события $(\mu_n = m)$ равна вероятности (2.13), умноженной на C_n^m, что снова приводит к формуле (2.12).

Формула (2.12) дает вероятность того, что при n-кратном повторении испытания случайное событие A произойдет точно m раз. Таким образом, мы получаем следующие таблицы распределения вероятностей для случайных величин μ_n и ω_n:

μ_n	n	$n-1$	...	m	...	1	0
	p^n	$np^{n-1}q$	...	$C_n^m p^m q^{n-m}$	...	npq^{n-1}	q^n

(2.14)

ω_n	1	$\frac{n-1}{n}$	...	$\frac{m}{n}$	...	$\frac{1}{n}$	0
	p^n	$np^{n-1}q$	...	$C_n^m p^m q^{n-m}$	...	npq^{n-1}	q^n

(2.15)

Распределение вероятностей, определяемое таблицей (2.14), называется *биномиальным распределением вероятностей*.

Пример. Из большой партии изделий берут на пробу 10 штук. Известно, что доля нестандартных изделий во всей партии составляет 25 %; требуется найти вероятность того, что более пяти отобранных изделий окажутся нестандартными.

Отбор каждого изделия будем считать испытанием, а обнаружение нестандартности у отобранного изделия — случайным событием A. Вероятность p события A равна, очевидно, доле нестандартных изделий во всей партии, то есть

$$p = 0{,}25.$$

Количество нестандартных изделий среди десяти отобранных есть случайная величина μ_{10} — частота повторения события A при десяти испытаниях. Задача сводится к вычислению вероятности того, что $\mu_{10} > 5$; по правилу сложения вероятностей

$$\mathbf{P}\{\mu_{10} > 5\} = \mathbf{P}\{\mu_{10} = 6\} + \mathbf{P}\{\mu_{10} = 7\} + \ldots + \mathbf{P}\{\mu_{10} = 10\}.$$

Подсчет вероятностей по биномиальной формуле (2.12) при $p=0{,}25$, $q=0{,}75$ и $n=10$ приводит к следующей таблице распределения (вероятности округлены до 0,0001):

Количество нестандартных изделий m	Вероятность $\mathbf{P}\{\mu_{10}=m\}$
0	0,0563
1	0,1877
2	0,2816
3	0,2503
4	0,1460
5	0,0584
6	0,0162
7	0,0031
8	0,0004
9	0,0000
10	0,0000

Отсюда видно, что $\mathbf{P}\{\mu_{10}>5\}\approx 0{,}020$. Эта вероятность довольно мала, так что если бы в отобранном десятке изделий оказалось шесть (или более) нестандартных, то мы могли бы усомниться в том, действительно ли доля нестандартных изделий во всей партии составляет только 25%.

П р и м е ч а н и е. Наше решение, опирающееся на биномиальное распределение вероятностей величины μ_{10}, будет точным лишь в том случае, когда отбор изделий для пробы производится по *схеме случайной повторной выборки* (см. сноску на стр. 36). Для пояснения разницы между повторной и бесповторной выборками рассмотрим пример. Пусть в урне лежит $N=100$ шаров, из них $M=25$ белых, так что $p=0{,}25$, и пусть из урны последовательно вынимаются 2 шара. Сравним распределение вероятностей числа вынутых белых шаров при повторной и бесповторной выборках. И в том и в другом случае число вынутых бе́лых шаров равно сумме двух характеристических случайных величин λ_1 и λ_2, где λ_k — число белых шаров, появляющихся при вынимании k-го шара; λ_1 и λ_2 имеют одно и то же распределение вероятностей (см. пример 1 § 5, стр. 27):

λ_1	1	0
	$\frac{25}{100}$	$\frac{75}{100}$

,

λ_2	1	0
	$\frac{25}{100}$	$\frac{75}{100}$

Но при повторной выборке величины λ_1 и λ_2 будут независимы, а при бесповторной выборке они уже будут зависимы. Поэтому распределение

вероятностей их суммы для повторной выборки дается таблицей

μ_2	2	1	0
	$\left(\frac{25}{100}\right)^2$	$2\cdot\frac{25}{100}\cdot\frac{75}{100}$	$\left(\frac{75}{100}\right)^2$

а для бесповторной выборки — таблицей

$\overset{*}{\mu}_2$	2	1	0
	$\frac{25}{100}\cdot\frac{24}{99}$	$\frac{25}{100}\cdot\frac{75}{99}+\frac{75}{100}\cdot\frac{25}{99}$	$\frac{75}{100}\cdot\frac{74}{99}$

Сравнивая таблицы распределения величин μ_2 и $\overset{*}{\mu}_2$, мы замечаем, что соответственные вероятности в них мало отличаются друг от друга. Очевидно, это отличие было бы еще меньше, если бы мы взяли, например, $N=1000$, $M=250$.

Подобные рассуждения позволяют заключить, что приведенное выше решение задачи, основанное на биномиальном распределении, будет тем более точным, чем меньше объем выборки по сравнению с числами M и N.

§ 8. Непрерывные случайные величины

Дискретные случайные величины не исчерпывают всех типов случайных величин. В теории вероятностей часто приходится вести расчет с такими случайными величинами, возможные значения которых сплошь заполняют некоторый интервал, как, например, упомянутые во введении отклонения размеров деталей от номинала. Такие случайные величины получили название непрерывных *).

Закон распределения вероятностей для непрерывной случайной величины ξ должен позволять находить вероятность попадания ее значения в любой интервал $(x_1;\ x_2)$; мы будем обозначать эту вероятность через $\mathbf{P}\{x_1<\xi<x_2\}$.

Пример. Равномерное распределение вероятностей.

В простейшем случае все возможные значения случайной величины ξ заполняют некоторый конечный интервал $(\alpha_1;\ \alpha_2)$

*) При некоторых условиях, которые будут уточнены далее.

и вероятность $\mathbf{P}\{x_1 < \xi < x_2\}$ для любого интервала $(x_1;\ x_2)$, лежащего внутри $(\alpha_1;\ \alpha_2)$, пропорциональна длине этого интервала:

$$\mathbf{P}\{x_1 < \xi < x_2\} = \lambda(x_2 - x_1) \qquad (\alpha_1 \leqslant x_1 < x_2 \leqslant \alpha_2). \quad (2.16)$$

Коэффициент λ должен быть выбран таким образом, чтобы имело место второе основное свойство вероятностей *); поскольку все возможные значения величины ξ лежат в $(\alpha_1;\ \alpha_2)$, то попадание значения ξ в этот интервал есть достоверное событие и поэтому его вероятность должна быть равна 1:

$$\mathbf{P}\{\alpha_1 < \xi < \alpha_2\} = \lambda(\alpha_2 - \alpha_1) = 1.$$

Отсюда однозначно определяется значение λ:

$$\lambda = \frac{1}{\alpha_2 - \alpha_1}.$$

Если распределение вероятностей величины ξ задается формулой (2.16), то говорят, что *случайная величина* ξ *равномерно распределена в интервале* $(\alpha_1;\ \alpha_2)$ или что величина ξ подчиняется закону равномерного распределения вероятностей.

Плотность распределения вероятностей

Если случайная величина ξ равномерно распределена в интервале $(\alpha_1;\ \alpha_2)$, то для любых точек x_1, x_2 этого интервала отношение

$$\frac{\mathbf{P}\{x_1 < \xi < x_2\}}{x_2 - x_1} \quad (2.17)$$

вероятности $\mathbf{P}\{x_1 < \xi < x_2\}$ к длине интервала $(x_1;\ x_2)$ есть величина постоянная (она равна $\lambda = \frac{1}{\alpha_2 - \alpha_1}$). Это отношение называется плотностью распределения вероятностей при равномерном распределении случайной величины ξ.

Для любой непрерывной случайной величины ξ отношение (2.17) уже может не быть постоянным. По аналогии с тем, как это делается в механике (при изучении распределения масс), мы вводим здесь понятие плотности распределения вероятностей в точке.

*) Первое основное свойство обеспечивается положительностью коэффициента λ, а третье свойство вытекает из того, что при объединении интервалов их длины складываются.

Плотностью распределения вероятностей случайной величины ξ *в точке* x *называется предел отношения*

$$\lim_{\Delta x \to 0} \frac{\mathbf{P}\{x < \xi < x + \Delta x\}}{\Delta x} = \varphi(x). \qquad (2.18)$$

Мы будем рассматривать лишь такие случайные величины, для которых этот предел существует в каждой точке x. Для таких величин вероятность попадания значения ξ в интервал $(x;\ x + dx)$ допускает выделение главной части, пропорциональной dx:

$$\mathbf{P}\{x < \xi < x + dx\} \approx \varphi(x)\,dx$$

(с точностью до малых высшего порядка относительно dx).

Эта главная часть называется *дифференциалом вероятности* и обозначается через $d\mathbf{P}_x$:

$$d\mathbf{P}_x = \varphi(x)\,dx. \qquad (2.19)$$

Зная дифференциал вероятности, мы можем с помощью интегрирования найти вероятность попадания значения ξ в любой интервал $(x_1;\ x_2)$:

$$\mathbf{P}\{x_1 < \xi < x_2\} = \int_{x_1}^{x_2} \varphi(x)\,dx. \qquad (2.20)$$

Таким образом, *для задания закона распределения непрерывной случайной величины достаточно задать плотность распределения ее вероятностей, то есть функцию* $\varphi(x)$ *). При всех расчетах с непрерывными случайными величинами дифференциал вероятности $\varphi(x)\,dx$ играет ту же роль, какую играют вероятности p_k при расчетах с дискретными случайными величинами: во многих формулах достаточно будет заменить p_k на $\varphi(x)\,dx$ и сумму — соответствующим интегралом, чтобы от формулы для дискретных величин перейти к формуле для непрерывных величин.

Замечание. Подчеркнем, что для непрерывной случайной величины ξ реальный смысл имеет только такое случайное событие, как попадание в интервал, а не попадание в отдельную точку. Поскольку вероятность попадания в малый интервал по определению должна быть

---

*) Строго говоря, непрерывная случайная величина ξ как раз и характеризуется тем, что вероятность попадания ее значения в любой интервал $(x_1;\ x_2)$ может быть представлена в виде интеграла (2.20) от некоторой функции $\varphi(x)$.

приближенно пропорциональна длине интервала, постольку *вероятность попадания в любую отдельную точку следует считать равной нулю.* Другими словами, если принятие непрерывной величиной ξ определенного значения рассматривать как случайное событие, то вероятность любого такого события должна быть равна нулю (хотя такое событие нельзя считать невозможным). На практике высказанное выше утверждение не приводит к недоразумениям, так как значение любой физической величины можно измерить лишь с некоторой точностью (абсолютно точное значение физической величины есть лишь математическая абстракция).

Основные свойства плотности распределения вероятностей

а) Плотность распределения $\varphi(x)$ неотрицательна для всех x. Это непосредственно следует из определения (2.18), где $\Delta x > 0$ и $\mathrm{P}\{x < \xi < x + \Delta x\} \geqslant 0$.

б) Интеграл от плотности распределения $\varphi(x)$, взятый по всему интервалу возможных значений случайной величины ξ, равен 1. Это следует из того, что указанный интеграл дает вероятность достоверного случайного события — принятия случайной величиной ξ какого-либо из своих значений. В зависимости от того, заполняют ли возможные значения величины ξ конечный интервал $(\alpha_1; \alpha_2)$ или всю числовую ось, рассматриваемое свойство записывается в виде

$$\int_{\alpha_1}^{\alpha_2} \varphi(x)\,dx = 1 \quad \text{или} \quad \int_{-\infty}^{\infty} \varphi(x)\,dx = 1.$$

Объединяя обе эти записи, пишут

$$\int \varphi(x)\,dx = 1, \tag{2.21}$$

не указывая пределов интегрирования; при этом подразумевается, что интеграл берется по всему интервалу возможных значений случайной величины ξ.

Заметим еще, что каждая неотрицательная функция $\varphi(x)$, удовлетворяющая условию (2.21), может служить плотностью распределения вероятностей некоторой случайной величины ξ.

Кривая распределения вероятностей. Кривой распределения вероятностей случайной величины называется график функции $y = \varphi(x)$, где $\varphi(x)$ — плотность распределения вероятностей (рис. 3). Кривая распределения может служить для графического расчета вероятностей, так как вероят-

ность $\mathbf{P}\{x_1 < \xi < x_2\}$ и площадь заштрихованной на рис. 3 криволинейной трапеции выражаются одним и тем же интегралом (2.20). Площадь заштрихованной криволинейной трапеции будет равна вероятности $\mathbf{P}\{x_1 < \xi < x_2\}$, если площадь всей фигуры, ограниченной кривой распределения и осью абсцисс, будет равна единице. Другими словами, *вероятность* $\mathbf{P}\{x_1 < \xi < x_2\}$ *равна отношению заштрихованной площади к площади всей фигуры* (принимаемой за единицу). Если же требуется выразить площадь сразу в безразмерных единицах, как и вероятность, то следует учитывать, что размерность $\varphi(x)$ равна $\dfrac{1}{\text{размерность } x}$.

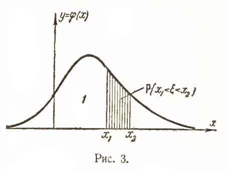


Рис. 3.

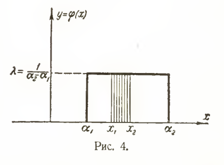


Рис. 4.

Для примера на рис. 4 дана кривая равномерного распределения вероятностей в интервале $(\alpha_1; \alpha_2)$.

Примеры непрерывных распределений вероятностей.

1) **Простейшее нормальное распределение вероятностей.** Говорят, что случайная величина ξ_0 имеет простейшее нормальное распределение вероятностей или что величина ξ_0 следует нормальному закону распределения, если ее плотность на всей числовой оси определена формулой

$$\varphi_0(x) = Ce^{-\frac{x^2}{2}}, \qquad \text{где} \qquad C = \frac{1}{\sqrt{2\pi}}. \qquad (2.22)$$

Значение коэффициента C выбрано так, чтобы выполнялось условие (2.21). Кривая распределения вероятностей дана на рис. 5. Она симметрична относительно оси ординат, при $x=0$ имеет максимум, равный $\frac{1}{\sqrt{2\pi}} \approx 0{,}4$, и имеет две точки перегиба при $x=\pm 1$. При $x \to \pm\infty$ кривая распределения асимптотически приближается к оси абсцисс, причем приближается весьма быстро (например, уже $\varphi_0(3)=0{,}0044$; $\varphi_0(4)=0{,}00013$).

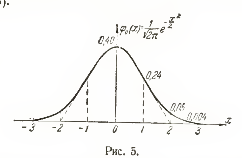


Рис. 5.

Нормальное распределение играет большую роль во многих приложениях теории вероятностей, в частности при обработке результатов измерений (см. главы V и VI). Так как интеграл от плотности $\varphi_0(x)$ не выражается в конечном виде через элементарные функции, то для расчета вероятностей составлены весьма подробные и достаточно точные таблицы специальной функции

$$\Phi(t)=\frac{2}{\sqrt{2\pi}}\int_0^t e^{-\frac{x^2}{2}}dx, \qquad (2.23)$$

называемой интегралом вероятностей. Краткая таблица интеграла вероятностей приведена в приложении. Функция $\Phi(t)$ является нечетной:

$$\Phi(-t)=\frac{2}{\sqrt{2\pi}}\int_0^{-t} e^{-\frac{x^2}{2}}dx=-\Phi(t);$$

поэтому в таблицах даются значения $\Phi(t)$ только для положительных значений t. При изменении t от 0 до $+\infty$ функция $\Phi(t)$ возрастает от 0 до 1, причем возрастает очень быстро: уже $\Phi(3)=0{,}9973$; $\Phi(4)=0{,}999937$. График функции $\Phi(t)$ дан на рис. 6.

С помощью функции $\Phi(t)$ можно вычислить вероятность попадания случайной величины ξ_0 в любой интервал $(x_1;\ x_2)$ следующим образом:

$$\mathbf{P}\{x_1 < \xi_0 < x_2\} = \int_{x_1}^{x_2} \varphi_0(x)\,dx = \frac{1}{\sqrt{2\pi}}\int_0^{x_2} e^{-\frac{x^2}{2}}dx - \frac{1}{\sqrt{2\pi}}\int_0^{x_1} e^{-\frac{x^2}{2}}dx,$$

откуда

$$\mathbf{P}\{x_1 < \xi_0 < x_2\} = \frac{1}{2}\Phi(x_2) - \frac{1}{2}\Phi(x_1). \qquad (2.24)$$

В частности, для симметричного интервала $(-t; +t)$ получаем:

$$\mathbf{P}\{-t < \xi_0 < +t\} = \frac{1}{2}\Phi(t) - \frac{1}{2}\Phi(-t) = \Phi(t). \quad (2.25)$$

Таким образом, при $t > 0$ функция $\Phi(t)$ дает вероятность попадания случайной величины ξ_0 в симметричный интервал $(-t; +t)$.

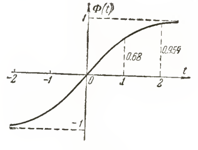


Рис. 6.

2) Общее нормальное распределение вероятностей. Так называется распределение с плотностью

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \qquad (2.26)$$

где $\sigma > 0$. При $a = 0$ и $\sigma = 1$ плотность $\varphi(x)$ превращается в плотность $\varphi_0(x)$ простейшего нормального распределения, рассмотренного в предыдущем примере. Кривые общего нормального распределения вероятностей при разных значениях σ (и $a = 0$) даны на рис. 7. Они отличаются от кривой простей-

шего нормального распределения (2.22) только изменением масштаба вдоль осей. С увеличением σ кривые распределения становятся более пологими. Кривая распределения при любом a отличается еще и сдвигом вдоль оси x (рис. 8); эта кривая распределения симметрична относительно прямой $x = a$. Расчет вероятностей в общем нормальном распределении производится с помощью интеграла вероятностей $\Phi(t)$, как будет показано в следующем параграфе (стр. 53).

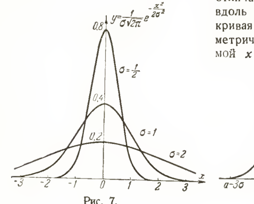


Рис. 7.

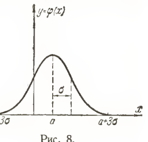


Рис. 8.

3) Примером несимметричного распределения вероятностей может служить распределение с плотностью

$$\varphi(x) = \begin{cases} 0 & \text{при } x \leqslant 0, \\ C_1 x^{\alpha-1} e^{-\beta x} & \text{при } x > 0 \quad (\alpha > 0,\ \beta > 0). \end{cases} \tag{2.27}$$

Коэффициент C_1 выбирается так, чтобы выполнялось условие (2.21):

$$C_1 = \frac{\beta^\alpha}{\Gamma(\alpha)},$$

где $\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1} e^{-x}\, dx$ есть гамма-функция Эйлера. Распределение (2.27) принадлежит к так называемым распределениям Пирсона; оно встречается во многих задачах гидроэнергетики. На рис. 9 приведена кривая распределения, отвечающая формуле (2.27) с $\alpha = \beta = 4$. Расчет вероятностей и здесь производится по специальным таблицам.

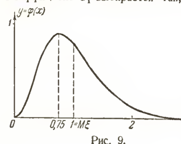


Рис. 9.

Функция распределения вероятностей

Функцией распределения вероятностей случайной величины ξ называется вероятность того, что величина ξ примет значение, меньшее некоторого числа x; эту функцию обозначим через $F(x) = \mathrm{P}\{\xi < x\}$.

Для дискретной случайной величины функция распределения равна сумме вероятностей тех ее значений x_k, которые меньше x: $F(x) = \sum\limits_{x_k < x} p_k$. Например, для случайной величины с таблицей распределения (2.3):

$$F(x) = \begin{cases} 0 & \text{при } x \leqslant 1, \\ 0{,}8 & \text{при } 1 < x \leqslant 2, \\ 0{,}96 & \text{при } 2 < x \leqslant 3, \\ 1 & \text{при } 3 < x. \end{cases}$$

Для непрерывной случайной величины в соответствии с формулой (2.20) функция распределения равна интегралу от плотности распределения:

$$F(x) = \int\limits_{-\infty}^{x} \varphi(t)\, dt.$$

Например, для простейшего нормального распределения (2.22) функция распределения выражается через интеграл вероятностей

$$F(x) = \int\limits_{-\infty}^{x} \varphi_0(t)\, dt = \frac{1}{2}\Phi(x) - \frac{1}{2}\Phi(-\infty) = \frac{1}{2}\Phi(x) + \frac{1}{2}.$$

Из определения $F(x)$ и основных свойств вероятности вытекает, что функция распределения вероятностей есть возрастающая функция,

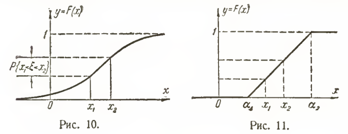


Рис. 10. Рис. 11.

изменяющаяся от 0 до 1. Ее график называют *интегральной кривой распределения вероятностей* (рис. 10). На рис. 11 дана, для примера, интегральная кривая равномерного распределения вероятностей в интервале $(\alpha_1;\ \alpha_2)$. Так как по правилу сложения вероятностей при $x_1 < x_2$

$$\mathrm{P}\{\xi < x_2\} = \mathrm{P}\{\xi < x_1\} + \mathrm{P}\{x_1 \leqslant \xi < x_2\},$$

то вероятность попадания значения ξ в интервал $(x_1;\ x_2)$ равна приращению функции распределения вероятностей

$$\mathbf{P}\{x_1 \leqslant \xi < x_2\} = F(x_2) - F(x_1).$$

Это позволяет графически находить вероятности по интегральной кривой распределения (см. рис. 10), если за единицу масштаба оси ординат принять $F(+\infty) = 1$.

Функция распределения вероятностей пригодна для задания закона распределения как дискретных, так и непрерывных случайных величин (а также и для задания случайных величин более сложной природы). Однако ее применение в общем случае требует специального математического аппарата (интеграла Стильтьеса), что выходит за рамки настоящего пособия.

§ 9. Функции от случайных величин

Пусть $f(x)$ — однозначная функция, определенная на совокупности всех возможных значений x величины ξ. Под функцией $f(\xi)$ случайной величины ξ понимают такую случайную величину η, которая принимает значение $y = f(x)$ каждый раз, когда величина ξ принимает значение x. Например, если случайная величина ξ есть диаметр обтачиваемого на станке валика, то площадь поперечного сечения валика есть случайная величина $\eta = \frac{\pi}{4}\xi^2$.

Нашей задачей является установление связи между законами распределения вероятностей случайных величин ξ и $\eta = f(\xi)$. Начнем с функции от дискретной случайной величины

ξ	x_1	x_2	...
	p_1	p_2	...

Если в результате испытания величина ξ примет некоторое значение x_k, то случайная величина $\eta = f(\xi)$ примет значение $f(x_k)$. Но вероятность, например, события $\eta = f(x_1)$ может быть больше вероятности p_1 события $\xi = x_1$, если среди значений $f(x_k)$ встретятся числа, равные $f(x_1)$. А именно, по правилу сложения вероятностей мы должны считать вероятность случайного события $\eta = f(x_1)$ равной сумме всех тех вероятностей p_k, для которых числа $f(x_k)$ равны $f(x_1)$.

На практике для построения таблицы распределения функции $f(\xi)$ обычно строят сначала вспомогательную таблицу

$f(x_1)$	$f(x_2)$	...
p_1	p_2	...

(2.28)

а затем объединяют в ней равные значения $f(x_k)$, складывая соответствующие вероятности. Заметим, что если все значения $f(x_k)$ различны, то таблица (2.28) будет готовой таблицей распределения функции $f(\xi)$ (см., например, (2.6)).

Примеры.

1) Рассмотрим степени λ^n $(n = 1, 2, 3, \ldots)$ от характеристической случайной величины λ с распределением вероятностей (2.9). Все эти степени имеют то же распределение, что и λ, так как $1^n = 1$, $0^n = 0$.

2) Рассмотрим функцию $\sin\left(\frac{\pi}{2}\xi\right)$ от случайной величины

ξ	1	2	3	...	n	...
	$\frac{1}{2}$	$\frac{1}{2^2}$	$\frac{1}{2^3}$	...	$\frac{1}{2^n}$	...

.

Так как

$$\sin\left(\frac{\pi}{2}n\right) = \begin{cases} 0 & \text{при четном } n, \\ 1 & \text{при } n = 4k+1, \\ -1 & \text{при } n = 4k+3, \end{cases}$$

то таблицей распределения для $\sin\left(\frac{\pi}{2}\xi\right)$ будет:

$\sin\left(\frac{\pi}{2}\xi\right)$	0	1	-1
	p_0	p_1	p_{-1}

,

где

$$p_0 = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \ldots = \frac{1}{4\left(1 - \frac{1}{4}\right)} = \frac{1}{3};$$

$$p_1 = \frac{1}{2} + \frac{1}{2^5} + \frac{1}{2^9} + \ldots = \frac{1}{2\left(1 - \frac{1}{16}\right)} = \frac{8}{15};$$

$$p_{-1} = \frac{1}{2^3} + \frac{1}{2^7} + \frac{1}{2^{11}} + \ldots = \frac{1}{8\left(1 - \frac{1}{16}\right)} = \frac{2}{15}.$$

3) Рассмотрим квадрат случайной величины

ξ	$-b$	b
	p	$1-p$

. Так как $(-b)^2 = b^2$, то получаем

ξ^2	b^2
	1

. Из последней таблицы распределения видно, что величина ξ^2 принимает единственное значение с вероятностью 1 и, значит, может рассматриваться как неслучайная величина.

Функция от непрерывной случайной величины ξ

Будем считать, что функция $f(x)$ непрерывна вместе с первой производной в интервале возможных значений x величины ξ. Нашей задачей является установление зависимости между плотностями распределения вероятностей $\varphi(x)$ и $\psi(y)$ случайных величин ξ и $\eta = f(\xi)$. Проще всего эта зависимость находится тогда, когда функция $f(x)$ *строго возрастает*. При этом каждый интервал $(x_1; x_2)$ отображается взаимно однозначно на соответствующий интервал $(y_1; y_2)$ (рис. 12)

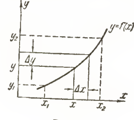


Рис. 12.

и поэтому вероятности попадания случайных величин ξ и η в соответственные интервалы должны быть равны. В применении к малым соответственным интервалам $(x;\ x+\Delta x)$ и $(y;\ y+\Delta y)$ это означает равенство дифференциалов вероятностей:

$$\varphi(x)\,dx = \psi(y)\,dy, \tag{2.29}$$

откуда и находят искомую зависимость:

$$\psi(y) = \varphi(x)\frac{dx}{dy} = \varphi[g(y)]\,g'(y), \tag{2.30}$$

где $x = g(y)$ есть функция, обратная функции $y = f(x)$.

Если функция $y = f(x)$ *строго убывает*, то положительному значению dx соответствует отрицательное значение dy. Поэтому в формуле (2.29) надо заменить dy на $-dy = |dy|$, что приводит к более

общей зависимости

$$\psi(y) = \varphi(x)\left|\frac{dx}{dy}\right| = \varphi[g(y)]\,|g'(y)|. \tag{2.31}$$

Пример. Линейная функция.

Для линейной функции $\eta = a + b\xi$ имеем: $y = f(x) = a + bx$; $x = g(y) = \frac{y-a}{b}$; $g'(y) = \frac{1}{b}$. Поэтому зависимость между плотностями распределения вероятностей величин ξ и η будет:

$$\psi(y) = \frac{1}{|b|}\varphi\left(\frac{y-a}{b}\right). \tag{2.32}$$

Например, если случайная величина ξ имеет равномерное распределение в интервале $(\alpha_1;\ \alpha_2)$, то случайная величина $\eta = a + b\xi$ будет иметь также равномерное распределение в интервале $(a + b\alpha_1;\ a + b\alpha_2)$; проверку этого мы предоставляем читателю.

Если случайная величина ξ_0 имеет простейшее нормальное распределение с плотностью (2.22), то величина $\eta = a + b\xi_0$ будет иметь общее нормальное распределение с плотностью

$$\psi(y) = \frac{1}{|b|\sqrt{2\pi}}e^{-\frac{(y-a)^2}{2b^2}}.$$

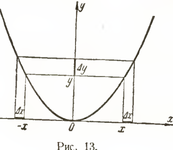


Рис. 13.

Это позволяет производить расчет вероятностей для общего нормального распределения (2.26) с помощью интеграла вероятностей (2.23). Действительно, пусть случайная величина ξ имеет общее нормальное распределение вероятностей с плотностью (2.26). Тогда случайная величина $\xi_0 = \frac{\xi - a}{\sigma}$ будет иметь простейшее нормальное распределение (2.22). При этом неравенство $x_1 < \xi < x_2$ равносильно неравенству

$$\frac{x_1 - a}{\sigma} < \xi_0 < \frac{x_2 - a}{\sigma},$$

что и приводит к искомой формуле:

$$\mathbf{P}\{x_1 < \xi < x_2\} = \mathbf{P}\left\{\frac{x_1 - a}{\sigma} < \xi_0 < \frac{x_2 - a}{\sigma}\right\} =$$

$$= \frac{1}{2}\Phi(t_2) - \frac{1}{2}\Phi(t_1), \tag{2.33}$$

где

$$t_1 = \frac{x_1 - a}{\sigma}; \quad t_2 = \frac{x_2 - a}{\sigma}.$$

Нахождение плотности распределения вероятностей для немонотонной функции мы покажем лишь на примере квадратичной функции $\eta = \xi^2$, причем будем считать, что величина ξ распределена на всей оси. Здесь $y = f(x) = x^2 \geqslant 0$; обратная функция имеет две однозначные ветви: $x = g_1(y) = +\sqrt{y}$; $x = g_2(y) = -\sqrt{y}$ (рис. 13). Применяя к каждой из этих ветвей формулу (2.31) и объединяя одинаковые значения y, получаем при $y > 0$:

$$\psi(y) = \varphi(\sqrt{y}) \left| \frac{1}{2\sqrt{y}} \right| + \varphi(-\sqrt{y}) \left| -\frac{1}{2\sqrt{y}} \right| = \\ = [\varphi(\sqrt{y}) + \varphi(-\sqrt{y})] \frac{1}{2\sqrt{y}}.$$

При $y < 0$ надо положить $\psi(y) = 0$.

Понятие о двумерных случайных величинах и функциях от двух случайных величин

Для изучения функций от нескольких случайных величин, а также для решения многих практических задач оказывается необходимым рассмотрение многомерных случайных величин, то есть величин, значения которых распределены в пространстве двух, трех и более измерений. Примером двумерной случайной величины может служить точка попадания в мишень. Если координаты этой точки в плоскости мишени обозначить через ξ и η, то мы получим двумерную случайную величину $(\xi; \eta)$.

Аналогичный пример дискретной двумерной величины рассмотрен выше, в примере 2 § 5 (стр. 27); приведенная там таблица (1.19) дает пример таблицы распределения вероятностей двумерной дискретной величины. В настоящей книге нет возможности подробно рассмотреть многомерные случайные величины. Мы укажем лишь некоторые формулы, относящиеся к непрерывным двумерным случайным величинам (соответствующие формулы для дискретных двумерных величин имеют аналогичный вид).

Значение величины $(\xi; \eta)$ есть точка $(x; y)$; распределение вероятностей задается *дифференциалом вероятности*

$$dP_{xy} = \varphi(x; y)\, dx\, dy, \tag{2.34}$$

дающим главную часть вероятности попадания точки $(\xi; \eta)$ в прямоугольник $\begin{bmatrix} x < \xi < x + dx \\ y < \eta < y + dy \end{bmatrix}$ (рис. 14).

Функция $\varphi(x, y)$ называется двумерной плотностью распределения вероятностей. Вероятность попадания точки $(\xi; \eta)$ в некоторую

область (D) определяется двойным интегралом:

$$\mathbf{P}\{(\xi;\ \eta)\in(D)\}=\iint\limits_{(D)}\varphi(x;\ y)\,dx\,dy. \tag{2.35}$$

Плотность $\varphi(x;y)$ может быть любой неотрицательной функцией, удовлетворяющей условию

$$\iint\varphi(x;\ y)\,dx\,dy=1,$$

где интеграл берется по всей области возможных значений случайной величины $(\xi;\ \eta)$.

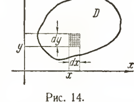


Рис. 14.

Простейшим примером непрерывной двумерной случайной величины является *случайная величина* $(\xi;\ \eta)$ *с равномерным распределением в некоторой конечной области* (D_0). Для такой величины вероятность попадания в любую область (D), лежащую внутри (D_0), пропорциональна площади S_D этой области. При этом

$$dP_{xy}=\lambda\,dx\,dy \quad \text{для точек внутри } (D_0),$$
$$dP_{xy}=0 \quad \text{для точек вне } (D_0).$$

Коэффициент пропорциональности λ находится из условия

$$\mathbf{P}\{(\xi;\ \eta)\in(D_0)\}=\lambda S_{D_0}=1,$$

откуда $\lambda=\dfrac{1}{S_{D_0}}$. Таким образом, вероятность попадания в область (D), лежащую внутри (D_0), равна отношению площадей S_D и S_{D_0}. Двумерная плотность здесь равна

$$\varphi(x;\ y)=\frac{1}{S_{D_0}}.$$

Координаты ξ и η двумерной непрерывной случайной величины будут одномерными непрерывными случайными величинами. Плотности $\psi_1(x)$ и $\psi_2(y)$ случайных величин ξ и η связаны с двумерной плотностью $\varphi(x;\ y)$ следующими формулами:

$$\psi_1(x)=\int\varphi(x;\ y)\,dy, \tag{2.36}$$

$$\psi_2(y)=\int\varphi(x;\ y)\,dx. \tag{2.37}$$

Для доказательства, например, формулы (2.36) достаточно заметить, что дифференциал вероятности $\psi_1(x)\,dx$ можно рассматривать

как вероятность попадания точки $(\xi;\ \eta)$ в заштрихованную на рис. 15 полосу и поэтому

$$\psi_1(x)\,dx = \left[\int \varphi(x;\ y)\,dy\right] dx.$$

Для дискретной величины с распределением (1.19) аналогичной формулой является формула (1.20).

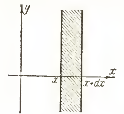


Рис. 15.

Случайные величины ξ *и* η *называются независимыми, если дифференциал вероятности* dP_{xy} *равен произведению дифференциалов вероятностей* $\psi_1(x)\,dx$ *и* $\psi_2(y)\,dy$, то есть если двумерная плотность $\varphi(x;\ y)$ равна произведению плотностей распределений величин ξ и η:

$$\varphi(x;\ y) = \psi_1(x)\,\psi_2(y). \tag{2.38}$$

Для дискретной величины с распределением (1.19) аналогичной формулой является формула (2.8).

Для функции $\zeta = f(\xi;\ \eta)$ *от двух случайных величин* ξ, η *распределение вероятностей определяется формулой*

$$\mathbf{P}\{z < \zeta < z + \Delta z\} = \iint\limits_{(D_{z,\Delta z})} \varphi(x;\ y)\,dx\,dy, \tag{2.39}$$

где $(D_{z,\Delta z})$ есть такая область плоскости $(x;\ y)$, в которой

$$z < f(x;\ y) < z + \Delta z,$$

а $\varphi(x;\ y)$ — плотность распределения двумерной величины $(\xi,\ \eta)$. Выделяя в интеграле (2.39) главную часть, линейную относительно Δz, находят дифференциал вероятности dP_z, а значит, и плотность распределения вероятностей функции $\zeta = f(\xi;\ \eta)$.

Пример. Распределение суммы случайных величин.

Для суммы $\zeta = \xi + \eta$ область $(D_{z,\Delta z})$ представляет собой полоску, заключенную между прямыми $x + y = z$ и $x + y = z + \Delta z$ (рис. 16). Поэтому

$$\mathbf{P}\{z < \zeta < z + \Delta z\} = \int\limits_{-\infty}^{\infty} dx \int\limits_{z-x}^{z+\Delta z - x} \varphi(x;\ y)\,dy.$$

Выделяя главную часть внутреннего интеграла:

$$\int\limits_{(z-x)}^{(z-x)+\Delta z} \varphi(x;\ y)\,dy \approx \varphi(x;\ z-x)\,\Delta z,$$

получаем дифференциал вероятности величины ζ:

$$dP_z = \int_{-\infty}^{\infty} \varphi(x; z - x)\, dx\, \Delta z,$$

откуда находим плотность $\chi(z)$ распределения суммы $\zeta = \xi + \eta$:

$$\chi(z) = \int_{-\infty}^{\infty} \varphi(x;\ z - x)\, dx. \quad (2.40)$$

Особый интерес представляет сложение независимых случайных величин ξ и η. В этом случае из формулы (2.38) вытекает, что плотность распределения суммы $\zeta = \xi + \eta$ выражается через плотности распределения слагаемых ξ и η по формуле

$$\chi(z) = \int_{-\infty}^{\infty} \psi_1(x)\, \psi_2(z - x)\, dx. \quad (2.41)$$

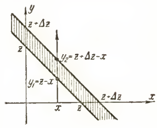


Рис. 16.

Интеграл $\int_{-\infty}^{\infty} \psi_1(x)\psi_2(z - x)\, dx$ называется *сверткой функций* ψ_1 и ψ_2 и обозначается через $\psi_1 * \psi_2$.

Упражнения

1. Найти распределение вероятностей суммы очков, выпадающих на трех правильных игральных костях. Проверить, что на трех костях вероятнее выбросить 11 очков, чем 12, хотя и 11 и 12 очков получаются при шести комбинациях:

11 очков при комбинациях $(6+4+1)$, $(6+3+2)$, $(5+5+1)$, $(5+4+2)$, $(5+3+3)$, $(4+4+3)$;
12 очков при комбинациях $(6+5+1)$, $(6+4+2)$, $(6+3+3)$, $(5+5+2)$, $(5+4+3)$, $(4+4+4)$.

2. Из урны, в которой лежит 20 черных и 4 белых шара, вынимаются 5 шаров. Найти распределение вероятностей числа ξ вынутых белых шаров.

Ответ.

ξ	0	1	2	3	4
p	$\frac{646}{1771}$	$\frac{1615}{3542}$	$\frac{285}{1771}$	$\frac{95}{5313}$	$\frac{5}{10626}$

3. Из урны, в которой лежит 20 черных и 4 белых шара, последовательно вынимаются шары до тех пор, пока не появится черный

шар. Найти распределение числа ξ вынутых при этом белых шаров, то есть числа белых шаров, вынутых до первого черного шара.

Ответ.

ξ	0	1	2	3	4
p	$\frac{5}{6}$	$\frac{10}{69}$	$\frac{5}{253}$	$\frac{10}{5313}$	$\frac{1}{10626}$

4. Найти сумму двух независимых величин с равномерным распределением вероятностей в интервале $(-1;\ +1)$.

Ответ. Плотность распределения вероятностей суммы

$$\varphi(x)=\begin{cases} 0 & \text{при } x<-2,\ x>2; \\ \frac{1}{4}(x+2) & \text{при } -2<x<0; \\ \frac{1}{4}(-x+2) & \text{при } 0<x<2. \end{cases}$$

5. Точка случайно попадает на окружность с равномерным распределением вероятностей по длине дуги. Найти распределение вероятностей проекции этой точки на диаметр.

Ответ. Плотность

$$\varphi(x)=\begin{cases} 0 & \text{при } x<-R,\ x>R, \\ \frac{1}{\pi\sqrt{R^2-x^2}} & \text{при } -R<x<R. \end{cases}$$

6. Кубики изготовляются с некоторой погрешностью. Считая, что линейные размеры кубиков имеют нормальное распределение вероятностей (2.26), найти распределение вероятностей их объемов (v).

Ответ. Плотность

$$\psi(v)=\frac{1}{3\sqrt[3]{v^2}\,\sigma\sqrt{2\pi}}e^{-\frac{\left(\sqrt[3]{v}-a\right)^2}{2\sigma^2}}.$$

7. Доказать, что наиболее вероятным значением частоты μ_n является целое число m_0, удовлетворяющее неравенству $np+p-1\leqslant m_0\leqslant np+p$; в частности, если число np — целое, то $m_0=np$.

Указание. Рассмотреть отношение

$$\frac{\mathbf{P}\{\mu_n=m+1\}}{\mathbf{P}\{\mu_n=m\}}=\frac{n-m}{m+1}\frac{p}{q}.$$

8. Задача о встрече. Два лица, A и B, условились встретиться в определенном месте между 0 и 1 часами. Пришедший пер-

вым ждет другого в течение 20 минут, после чего уходит. Определить вероятность p встречи лиц A и B, если моменты их прихода независимы и равномерно распределены в интервале $(0; 1)$.

Решение. Обозначим моменты прихода лиц A и B соответственно через ξ и η. По условию ξ и η независимы и равномерно распределены в интервале $(0; 1)$; поэтому случайная точка $(\xi; \eta)$ равномерно распределена в квадрате со стороной 1 (рис. 17). Задача заключается в нахождении вероятности неравенства $|\eta-\xi| \leqslant \frac{1}{3}$, то есть вероятности попадания точки $(\xi; \eta)$ в заштрихованную на рис. 17 полосу между прямыми $y-x=\frac{1}{3}$ и $y-x=-\frac{1}{3}$. Эта вероятность равна отношению заштрихованной площади к площади всего квадрата, то есть

Рис. 17.

$$p=\frac{1-\left(\frac{2}{3}\right)^2}{1}=\frac{5}{9}.$$

ГЛАВА III

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

При расчетах с дискретными или непрерывными случайными величинами не всегда целесообразно пользоваться таблицами или плотностями распределения вероятностей. Не говоря уже о том, что таблицы или плотности распределения не всегда бывают точно известны, расчеты с ними часто бывают сложны или громоздки. Оказывается, что ряд практически важных задач можно решить с помощью немногих осредненных характеристик распределения. Изучим сначала саму операцию осреднения, позволяющую получать подобные характеристики.

§ 10. Осреднение. Математическое ожидание случайной величины

Начнем с наиболее простого понятия среднего арифметического значения. Пусть имеется совокупность N элементов, различающихся величиной некоторого признака x (например, партия электроламп, различающихся сроком службы; или совокупность дождливых дней в году, различающихся величиной осадков в данном месте).

Средним арифметическим значением признака x в совокупности называется отношение суммы значений признака x у всех элементов совокупности к общему числу этих элементов.

Обозначим через $x_1, x_2, \ldots, x_\nu$ различные значения рассматриваемого признака у элементов совокупности; через M_k — количество элементов, у которых значение признака равно $x_k (k = 1, 2, \ldots, \nu)$; через $N = M_1 + M_2 + \ldots + M_\nu$ — общее число элементов совокупности. Тогда среднее арифметическое значение $\overline{x}$ представится формулой

$$\overline{x} = \frac{x_1M_1 + x_2M_2 + \ldots + x_\nu M_\nu}{N}.$$

Запишем эту формулу в виде

$$\bar{x} = x_1 \frac{M_1}{N} + x_2 \frac{M_2}{N} + \ldots + x_\nu \frac{M_\nu}{N}. \quad (3.1)$$

Из последней формулы видно, что среднее арифметическое значение зависит не от абсолютных количеств $M_1, M_2, \ldots, M_\nu$, а только от относительных количеств $\frac{M_1}{N}, \frac{M_2}{N}, \ldots, \frac{M_\nu}{N}$.

Переходя теперь к случайным величинам, рассмотрим сначала дискретную величину специального вида. Пусть из указанной выше совокупности выбирается наудачу один элемент *). Величина признака у выбираемого элемента есть дискретная случайная величина ξ со следующей таблицей распределения вероятностей:

ξ	x_1	x_2	...	x_k	...	x_ν
	$\frac{M_1}{N}$	$\frac{M_2}{N}$	...	$\frac{M_k}{N}$	...	$\frac{M_\nu}{N}$

(3.2)

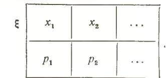
(так как величина ξ может принимать только те значения $x_1, x_2, \ldots, x_\nu$, которые имеются у элементов совокупности, а вероятность выбора элемента со значением признака x_k, очевидно, равна $\frac{M_k}{N}$). Среднее арифметическое значение признака в совокупности играет здесь роль среднего «ожидаемого» значения случайной величины ξ. Это среднее значение называется «математическим ожиданием» случайной величины ξ и обозначается через $\mathbf{M}\xi$. Таким образом, математическое ожидание случайной величины (3.2) равно

$$\mathbf{M}\xi = x_1 \frac{M_1}{N} + x_2 \frac{M_2}{N} + \ldots + x_\nu \frac{M_\nu}{N};$$

но здесь уже эту сумму надо толковать как сумму произведений значений величины ξ на их вероятности. Это позволяет

*) Так же, как и в § 2, проще всего представить себе это испытание как вынимание наугад шара из урны, где находится N одинаковых на ощупь шаров, из которых M_1 шаров имеют метку «x_1», M_2 шаров имеют метку «x_2», и т. д., причем перед выниманием все шары тщательно перемешивают.

сразу же распространить понятие математического ожидания на любую дискретную случайную величину



ξ	x_1	x_2	...
	p_1	p_2	...

Определение 1. *Математическим ожиданием* $\mathbf{M}\xi$ *дискретной случайной величины* ξ *называется сумма произведений всех ее возможных значений* (x_k) *на их вероятности* (p_k):

$$\mathbf{M}\xi = x_1p_1 + x_2p_2 + \ldots,$$

или, короче,

$$\mathbf{M}\xi = \sum x_k p_k, \tag{3.3}$$

где сумма берется по всем возможным значениям случайной величины ξ. Если множество этих возможных значений бесконечно, то мы будем предполагать еще, что бесконечный ряд (3.3) сходится абсолютно (в противном случае говорят, что математическое ожидание $\mathbf{M}\xi$ не существует; таких случайных величин мы не будем рассматривать).

Теперь мы можем распространить понятие математического ожидания на непрерывные случайные величины, учитывая, что для них роль вероятности p_k играет дифференциал вероятности $dP_x = \varphi(x)\,dx$.

Определение 2. *Математическим ожиданием* $\mathbf{M}\xi$ *непрерывной случайной величины* ξ *называется интеграл от произведения ее значений* x *на плотность распределения вероятностей* $\varphi(x)$:

$$\mathbf{M}\xi = \int x\varphi(x)\,dx, \tag{3.4}$$

причем интеграл (3.4) берется по всему интервалу возможных значений величины ξ. Этот интеграл часто записывают в виде $\int\limits_{-\infty}^{\infty} x\varphi(x)\,dx$ даже в том случае, когда возможные значения величины ξ заполняют конечный интервал; в этом случае полагают $\varphi(x) = 0$ вне указанного интервала. Если же возможные значения величины ξ заполняют бесконечный интервал, то мы будем предполагать, что несобственный интеграл

$\int\limits_{-\infty}^{\infty} x\varphi(x)\,dx$ сходится абсолютно (в противном случае говорят, что математическое ожидание $\mathbf{M}\xi$ не существует; таких величин мы не будем рассматривать).

Важно отметить, что все свойства математического ожидания (или, точнее, свойства самой операции осреднения) совершенно одинаковы как для дискретных, так и для непрерывных случайных величин *).

Свойства математического ожидания

Важнейшим свойством операции осреднения является линейность: *математическое ожидание линейной комбинации случайных величин равно линейной комбинации их математических ожиданий:*

$$\mathbf{M}(C_1\xi_1 + C_2\xi_2 + \ldots + C_n\xi_n) = C_1\mathbf{M}\xi_1 + C_2\mathbf{M}\xi_2 + \ldots + C_n\mathbf{M}\xi_n, \tag{3.5}$$

где $C_1, C_2, \ldots, C_n$ — постоянные.

Для доказательства свойства линейности достаточно доказать следующие теоремы.

1. Постоянный множитель C можно выносить за знак математического ожидания:

$$\mathbf{M}C\xi = C\mathbf{M}\xi. \tag{3.6}$$

2. Математическое ожидание суммы двух случайных величин равно сумме их математических ожиданий *(теорема сложения математических ожиданий):*

$$\mathbf{M}(\xi + \eta) = \mathbf{M}\xi + \mathbf{M}\eta. \tag{3.7}$$

Формула (3.6) особенно просто доказывается для дискретной величины ξ, умножение которой на постоянную C определено таблицей (2.6):

$$\mathbf{M}C\xi = \sum (Cx_k)\,p_k = C\sum x_k p_k = C\mathbf{M}\xi.$$

Для непрерывных величин доказательство дано ниже (стр. 66).

Теорему сложения (3.7) докажем для непрерывных случайных величин ξ, η. Обозначим через $\varphi(x; y)$ плотность

*) Можно дать единое определение математического ожидания для любой случайной величины, но это требует знакомства с аппаратом интеграла Стильтьеса, что выходит за рамки настоящего пособия.

совместного распределения, а через $\chi(z)$ — плотность распределения их суммы $\zeta = \xi + \eta$. Тогда по формулам (3.4) и (2.40) имеем:

$$M\zeta = \int z\chi(z)\,dz = \int_{-\infty}^{\infty} z\,dz \int_{-\infty}^{\infty} \varphi(x;\ z-x)\,dx.$$

Переменив порядок интегрирования, заменим z на $x+y$:

$$M(\xi+\eta) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} z\varphi(x;\ z-x)\,dz = \\ = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} (x+y)\,\varphi(x;\ y)\,dy. \qquad (3.8)$$

Воспользуемся теперь линейностью интеграла и формулами (2.36) и (2.37):

$$M(\xi+\eta) = \int_{-\infty}^{\infty} x\,dx \int_{-\infty}^{\infty} \varphi(x;\ y)\,dy + \int_{-\infty}^{\infty} y\,dy \int_{-\infty}^{\infty} \varphi(x;\ y)\,dx = \\ = \int_{-\infty}^{\infty} x\psi_1(x)\,dx + \int_{-\infty}^{\infty} y\psi_2(y)\,dy = M\xi + M\eta.$$

Для дискретных величин доказательство проводится тем же способом. Прежде всего, с помощью вспомогательной таблицы (2.7) легко убедиться в том, что математическое ожидание суммы $\xi + \eta$ может быть подсчитано по формуле

$$M(\xi+\eta) = \sum_{k,\,l} (x_k + y_l)\,p_{kl},$$

где сумма берется по всем значениям x_k и y_l. Далее, разлагая эту сумму на две:

$$\sum_{k,\,l} (x_k + y_l)\,p_{kl} = \sum_{k,\,l} x_k p_{kl} + \sum_{k,\,l} y_l p_{kl}$$

и преобразуя первую из них с помощью формулы (1.20), получаем:

$$\sum_{k,\,l} x_k p_{kl} = \sum_k x_k (p_{k1} + p_{k2} + \ldots) = \sum_k x_k p_k,$$

где $p_k = P\{\xi = x_k\}$. Таким образом, $\sum_{k,\,l} x_k p_{kl} = M\xi$, и аналогично $\sum_{k,\,l} y_l p_{kl} = M\eta$, что и доказывает формулу (3.7).

Отметим еще два свойства математического ожидания.

3. Математическое ожидание постоянной (неслучайной) величины C равно самой этой величине C.

Действительно, постоянную C можно рассматривать как случайную величину с единственным возможным значением C, вероятность которого равна 1. Поэтому $MC = C \cdot 1 = C$.

4. Математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий *(теорема умножения математических ожиданий):*

$$M\xi\eta = M\xi M\eta \text{ для независимых } \xi, \ \eta. \qquad (3.9)$$

Проведем доказательство для дискретных величин. В силу независимости величин ξ и η распределение вероятностей их произведения $(\xi\eta)$ определяется с помощью таблицы

x_1y_1	x_1y_2	x_2y_1	x_2y_2	...
p_1q_1	p_1q_2	p_2q_1	p_2q_2	...

,

в которой при необходимости следует объединить равные числа в первой строке, сложив соответствующие вероятности (ср. определение суммы дискретных величин на стр. 33). Поэтому математическое ожидание произведения $\xi\eta$ можно записать в виде

$$M\xi\eta = \sum_{k,\,l} (x_k y_l)(p_k q_l),$$

где сумма берется по всем возможным значениям x_k и y_l величин ξ и η соответственно. Произведя здесь внутреннее суммирование по l, можно вынести за знак внутренней суммы множитель $(x_k p_k)$, не зависящий от l:

$$M\xi\eta = \sum_k \sum_l x_k y_l p_k q_l = \sum_k x_k p_k \sum_l y_l q_l,$$

что и приводит к формуле (3.9).

Доказательство этой формулы для непрерывных случайных величин легко проводится с помощью приводимой далее формулы (3.13), что мы предоставляем сделать читателю.

Формула (3.9) без труда обобщается на любое число взаимно независимых сомножителей.

Правила вычисления математического ожидания от функции

1. Пусть ξ — дискретная случайная величина, принимающая значения x_k с вероятностью p_k. Функция $f(\xi)$ есть снова дискретная величина, и ее математическое ожидание определяется формулой

$$\mathbf{M}f(\xi) = \sum f(x_k)\,\mathbf{P}\{f(\xi) = f(x_k)\}, \tag{3.10}$$

где сумма берется по всем различным значениям $f(x_k)$.

Оказывается, что математическое ожидание функции $f(\xi)$ может быть вычислено без нахождения распределения вероятностей этой функции непосредственно по распределению самой величины ξ. А именно, имеет место формула

$$\mathbf{M}f(\xi) = \sum f(x_k)\,p_k, \tag{3.11}$$

где сумма берется по всем значениям x_k величины ξ.

Прежде чем доказывать формулу (3. 11) в общем случае, заметим, что если все значения $f(x_k)$ различны, то функция $f(\xi)$ имеет таблицу распределения (2.28), и в этом случае формула (3.11) полностью совпадает с формулой (3.10). В общем случае среди значений $f(x_k)$ могут встретиться равные числа; пусть, для определенности, равны только два значения: $f(x_1) = f(x_2)$. Тогда вероятность события $f(\xi) = f(x_1)$ равна $p_1 + p_2$, и поэтому соответствующее слагаемое в формуле (3.10) можно преобразовать так:

$$f(x_1)\,\mathbf{P}\{f(\xi) = f(x_1)\} = f(x_1)(p_1 + p_2) = f(x_1)\,p_1 + f(x_2)\,p_2,$$

что снова приводит к формуле (3.11).

2. Математическое ожидание функции $f(\xi)$ от непрерывной величины ξ тоже может быть вычислено непосредственно по плотности $\varphi(x)$ распределения самой величины ξ с помощью формулы

$$\mathbf{M}f(\xi) = \int f(x)\,\varphi(x)\,dx. \tag{3.12}$$

Мы ограничимся доказательством этой формулы только для случая возрастающей функции $f(x)$. Обозначим плотность распределения величины $\eta = f(\xi)$ через $\psi(y)$ и в формуле математического ожидания

$$\mathbf{M}\eta = \int y\psi(y)\,dy$$

сделаем замену переменной интегрирования $y = f(x)$; при этом по формуле (2.29) мы будем иметь $\psi(y)\,dy = \varphi(x)\,dx$, что сразу приводит к формуле (3.12). Например, для функции $C\xi$ формула (3.12) дает:

$$\mathbf{M}C\xi = \int Cx\varphi(x)\,dx = C\int x\varphi(x)\,dx = C\mathbf{M}\xi,$$

что доказывает формулу (3.6) для непрерывных величин.

3. Приведем уже без доказательства соответствующие правила вычисления математического ожидания от функции двух переменных.

Для дискретных величин

$$\mathbf{M}f(\xi;\ \eta) = \sum_{k,\,l} f(x_k;\ y_l)\,p_{kl},$$

где сумма берется по всем значениям x_k и y_l величин ξ и η, а p_{kl} есть вероятность совмещения случайных событий $(\xi = x_k)$ и $(\eta = y_l)$.

Для непрерывных величин

$$\mathbf{M}f(\xi;\ \eta) = \iint f(x;\ y)\,\varphi(x;\ y)\,dx\,dy, \tag{3.13}$$

где $\varphi(x;\ y)$ — плотность распределения случайной точки $(\xi;\ \eta)$.

Отметим, что частные случаи этих формул при $f(x;\ y) = x + y$ и $f(x;\ y) = xy$ мы уже рассмотрели ранее (см., например, формулу (3.8)).

§ 11. Центр распределения случайной величины

Математическое ожидание случайной величины дает удобную числовую характеристику ее расположения. Имея ту же размерность, что и значения случайной величины, математическое ожидание находится внутри интервала возможных ее значений; например, если все значения случайной величины ξ лежат в интервале $(\alpha_1;\ \alpha_2)$, то

$$\mathbf{P}\{\alpha_1 < \xi < \alpha_2\} = \int_{\alpha_1}^{\alpha_2} \varphi(x)\,dx = 1$$

и из неравенства

$$\int_{\alpha_1}^{\alpha_2} \alpha_1 \varphi(x)\,dx < \int_{\alpha_1}^{\alpha_2} x\varphi(x)\,dx < \int_{\alpha_1}^{\alpha_2} \alpha_2 \varphi(x)\,dx$$

следует, что

$$\alpha_1 < \mathbf{M}\xi < \alpha_2\ ^{*)}.$$

В следующей главе будет показано еще, что около математического ожидания случайной величины группируются средние арифметические из ее опытных значений. Чтобы подчеркнуть роль математического ожидания в качестве основной характеристики расположения случайной величины (в отличие от самой операции осреднения), мы дадим следующее **определение**:

---

*) В частности, если все значения $\xi > 0$, то и $\mathbf{M}\xi > 0$.

Центром распределения вероятностей случайной величины ξ называется ее математическое ожидание $M\xi$ *).

Поясним на примерах понятие центра распределения как числовой характеристики расположения.

1. Пусть ξ — число израсходованных патронов при стрельбе по схеме примера 2 § 2 (стр. 30). По таблице распределения (2.3) находим математическое ожидание числа израсходованных патронов:

$$M\xi = 1\cdot 0{,}8 + 2\cdot 0{,}16 + 3\cdot 0{,}04 = 1{,}24.$$

Оно оказывается нецелым. Для того чтобы пояснить, какая практическая польза может быть от нашего подсчета, представим себе, что производится 100 стрельб по указанной схеме. Пусть ξ_k — число израсходованных патронов при k-й стрельбе; тогда

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_{100}$$

есть общее число израсходованных патронов при ста стрельбах. Подсчитаем его математическое ожидание, пользуясь свойством линейности и учитывая, что $M\xi_k = 1{,}24$ ($k = 1, 2, \ldots, 100$):

$$M\xi = M\xi_1 + M\xi_2 + \ldots + M\xi_{100} = 100\cdot 1{,}24 = 124.$$

Практически это означает, что на 100 подобных стрельб будет израсходовано в среднем 124 патрона.

2. *Центр распределения Пуассона* (2.5):

$$M\xi = \sum_{m=0}^{\infty} m\frac{a^m}{m!}e^{-a} = e^{-a}a\left(1 + a + \frac{a^2}{2!} + \ldots + \frac{a^{m-1}}{(m-1)!} + \ldots\right) = a.$$

Таким образом, выясняется смысл постоянной a в распределении Пуассона: число a есть математическое ожидание

*) Для пояснения термина «центр распределения вероятностей» укажем на механическую аналогию его с понятием центра распределения масс (центра тяжести): если, например, в точках $x_1, x_2, \ldots, x_\nu$ оси x сосредоточены массы $p_1, p_2, \ldots, p_\nu$, то центр тяжести x_c этой системы находится по формуле

$$x_c = \frac{x_1p_1 + x_2p_2 + \ldots + x_\nu p_\nu}{p_1 + p_2 + \ldots + p_\nu}.$$

При условии $p_1 + p_2 + \ldots + p_\nu = 1$ эта формула совпадает с формулой (3.3).

случайной величины ξ, подчиняющейся закону распределения Пуассона (2.5).

3. *Центр распределения частоты μ_n и относительной частоты ω_n случайного события.* Непосредственный подсчет математического ожидания по таблице распределения (2.14) приводит к формуле

$$M\mu_n = \sum_{m=0}^{n} mC_n^m p^m q^{n-m}.$$

Для более быстрого подсчета мы воспользуемся свойством линейности математического ожидания и представлениями (2.10) и (2.11) случайных величин μ_n и ω_n через характеристические случайные величины $\lambda_1, \lambda_2, \ldots, \lambda_n$. Прежде всего, непосредственный подсчет по таблицам распределения (2.9) дает:

$$M\lambda_k = 1p + 0q = p \qquad (k = 1, 2, \ldots, n). \qquad (3.14)$$

Это значит, что *центром распределения характеристической случайной величины служит вероятность случайного события*. Далее находим:

$$M\mu_n = M\lambda_1 + M\lambda_2 + \ldots + M\lambda_n = np; \qquad (3.15)$$

$$M\omega_n = \frac{1}{n} M\mu_n = p. \qquad (3.16)$$

Таким образом, *центром распределения относительной частоты (ω_n) случайного события является вероятность этого события при единичном испытании, а центр распределения частоты (μ_n) больше вероятности в n раз.* Заметим, что это вполне согласуется с нашим интуитивным представлением о математическом ожидании. Если, например, вероятность случайного события равна $p = 0{,}2$ и испытание повторяется $n = 100$ раз, то мы ожидаем, что случайное событие появится $np = 20$ раз. Точно так же, если нам говорят, что вероятность брака в большой партии изделий составляет $p = 1\% = 0{,}01$, то, проверяя на выборку $n = 1000$ изделий, мы склонны ожидать, что обнаружим $np = 10$ бракованных (конечно, мы допускаем возможность и некоторых отклонений, но здесь речь идет как раз о среднем ожидаемом результате). Заметим еще, что линейность математического ожидания позволяет нам из формулы (3.14) получить более общий результат, чем (3.16). Если случайное событие A в каждом

k-м испытании имеет свою вероятность p_k, то центр распределения относительной частоты ω_n события A при n испытаниях будет равен

$$\mathbf{M}\omega_n = \frac{1}{n}(\mathbf{M}\lambda_1 + \mathbf{M}\lambda_2 + \ldots + \mathbf{M}\lambda_n) = \frac{p_1 + p_2 + \ldots + p_n}{n},$$

то есть он будет равен среднему арифметическому из вероятностей события A во всех n испытаниях.

4. Если случайная величина ξ равномерно распределена на интервале $(\alpha_1; \alpha_2)$, то центр ее распределения совпадает с серединой этого интервала. Действительно, плотность равномерного распределения постоянна и равна $\frac{1}{\alpha_2 - \alpha_1}$ в интервале $(\alpha_1; \alpha_2)$, так что

$$\mathbf{M}\xi = \int_{\alpha_1}^{\alpha_2} x \frac{1}{\alpha_2 - \alpha_1} dx = \frac{1}{\alpha_2 - \alpha_1} \frac{\alpha_2^2 - \alpha_1^2}{2} = \frac{\alpha_1 + \alpha_2}{2}.$$

5. *Центр нормального распределения.* Для простейшего нормального распределения (2.22) центр распределения равен нулю, так как плотность $\varphi_0(x)$ есть четная функция

$$\mathbf{M}\xi_0 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} dx = 0.$$

Случайная величина ξ с общим нормальным распределением (2.26) может быть выражена через случайную величину ξ_0, как было показано на стр. 53:

$$\xi = a + \sigma\xi_0.$$

Поэтому

$$\mathbf{M}\xi = a + \sigma\mathbf{M}\xi_0 = a.$$

Таким образом, *центр общего нормального распределения равен его параметру a* (что позволяет выяснить смысл этого параметра и вполне согласуется с симметрией кривой нормального распределения относительно прямой $x = a$).

Замечание. Если кривая распределения вероятностей симметрична относительно некоторой прямой $x = a$, то центр распределения всегда совпадает с точкой a.

§ 12. Характеристики рассеяния случайной величины. Понятие о моментах распределеиия

Рассеяние случайной величины ξ связано с отклонением $\xi - a$ этой величины от ее центра распределения $a = \mathbf{M}\xi$. Непосредственное осреднение этого отклонения не может дать числовой характеристики рассеяния, так как

$$\mathbf{M}(\xi - a) = \mathbf{M}\xi - a = 0,$$

то есть отклонения противоположных знаков в среднем взаимно погашаются.

Основной числовой характеристикой рассеяния случайной величины ξ *является среднее квадратическое отклонение* σ, *определяемое по формуле*

$$\sigma = \sigma(\xi) = \sqrt{\mathbf{M}(\xi - a)^2}, \quad \text{где} \quad a = \mathbf{M}\xi. \qquad (3.17)$$

Стоящее под корнем среднее значение квадрата отклонения $\mathbf{M}(\xi - a)^2$ носит специальное название *дисперсии случайной величины* ξ и обозначается через $\mathbf{D}\xi$:

$$\mathbf{D}\xi = \mathbf{M}(\xi - a)^2 = \sigma^2(\xi).$$

Пользуясь формулами (3.11) и (3.12), запишем формулы для дисперсий дискретных и непрерывных случайных величин в виде *)

$$\sigma^2(\xi) = \sum (x_k - a)^2 p_k;$$
$$\sigma^2(\xi) = \int (x - a)^2 \varphi(x)\, dx.$$

Из написанных формул видно, что среднее квадратическое отклонение имеет ту же размерность, что и значения случайной величины. Особая роль среднего квадратического отклонения в оценке рассеяния случайной величины будет подробно выяснена дальше (главным образом в главах IV и V). В частности, дальше будет показано, что практически не встречаются такие значения случайной величины, отклонения

*) Здесь уместно продолжить механическую аналогию, указанную в сноске на стр. 68. Если толковать $p_1, p_2, \ldots, p_\nu$ как массы, сосредоточенные в точках $x_1, x_2, \ldots, x_\nu$ оси x, то дисперсию $\sigma^2 = \sum_{k=1}^{\nu} (x_k - a)^2 p_k$ можно рассматривать как центральный момент инерции этой системы материальных точек (относительно центра тяжести a).

которых от ее центра распределения во много раз больше, чем σ. Здесь же мы ограничимся рассмотрением примеров и простейших свойств среднего квадратического отклонения.

Основные правила вычислений средних квадратических отклонений и дисперсий

1) Если ξ — случайная величина, а C — постоянная, то

$$\sigma(C\xi) = |C|\,\sigma(\xi); \tag{3.18}$$

$$\sigma(\xi + C) = \sigma(\xi). \tag{3.19}$$

Эти формулы доказываются непосредственным подсчетом дисперсий:

$$\sigma^2(C\xi) = M(C\xi - MC\xi)^2 = \\ = M(C\xi - Ca)^2 = C^2 M(\xi - a)^2 = C^2\sigma^2(\xi);$$

$$\sigma^2(\xi + C) = M[(\xi + C) - M(\xi + C)]^2 = \\ = M[(\xi + C) - (a + C)]^2 = M(\xi - a)^2 = \sigma^2(\xi).$$

2) Если ξ и η — *независимые* случайные величины, то дисперсия их суммы равна сумме их дисперсий

$$\sigma^2(\xi + \eta) = \sigma^2(\xi) + \sigma^2(\eta) \tag{3.20}$$

(теорема сложения дисперсий) и, следовательно,

$$\sigma(\xi + \eta) = \sqrt{\sigma^2(\xi) + \sigma^2(\eta)}.$$

Доказательство формулы (3.20). Обозначим $M\xi = a$; $M\eta = b$; тогда $M(\xi + \eta) = a + b$, и поэтому

$$\sigma^2(\xi + \eta) = M[(\xi + \eta) - (a + b)]^2 = \\ = M[(\xi - a)^2 + 2(\xi - a)(\eta - b) + (\eta - b)^2].$$

В силу линейности математического ожидания имеем:

$$\sigma^2(\xi + \eta) = M(\xi - a)^2 + 2M(\xi - a)(\eta - b) + M(\eta - b)^2.$$

Так как по условию случайные величины ξ и η независимы, то можно применить теорему умножения математических ожиданий:

$$M(\xi - a)(\eta - b) = M(\xi - a)\,M(\eta - b).$$

Но, как было показано выше, $M(\xi - a) = 0$, и поэтому

$$M(\xi - a)(\eta - b) = 0,$$

что и приводит к формуле (3.20).

Теорема сложения дисперсий без труда обобщается на любое число попарно независимых случайных величин.

Следствие. Дисперсия линейной комбинации попарно независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ может быть вычислена по формуле

$$\sigma^2(C_1\xi_1 + C_2\xi_2 + \ldots + C_n\xi_n) = \\ = C_1^2\sigma^2(\xi_1) + C_2^2\sigma^2(\xi_2) + \ldots + C_n^2\sigma^2(\xi_n).$$

Это непосредственно вытекает из формул (3.20) и (3.18).

В частности, если все величины $\xi_1, \xi_2, \ldots, \xi_n$ имеют одну и ту же дисперсию

$$\sigma^2(\xi_k) = \sigma^2 \quad (k = 1, 2, \ldots, n),$$

то дисперсия их средней арифметической равна

$$\sigma^2\left(\frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n}\right) = \\ = \frac{1}{n^2}[\sigma^2(\xi_1) + \sigma^2(\xi_2) + \ldots + \sigma^2(\xi_n)] = \frac{\sigma^2}{n}.$$

Следовательно, среднее квадратическое отклонение ее равно

$$\sigma\left(\frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n}\right) = \frac{\sigma}{\sqrt{n}}. \qquad (3.21)$$

Последняя формула играет большую роль при обработке результатов измерений (см. далее главу VI).

Примеры.

1. *Среднее квадратическое отклонение относительной частоты.* Как показывает формула (2.11), относительная частота ω_n есть среднее арифметическое из взаимно независимых характеристических случайных величин $\lambda_1, \lambda_2, \ldots, \lambda_n$ с одинаковыми таблицами распределения (2.9):

$$\omega_n = \frac{\lambda_1 + \lambda_2 + \ldots + \lambda_n}{n}, \quad \text{где} \quad \lambda_k \begin{array}{|c|c|} \hline 1 & 0 \\ \hline p & q \\ \hline \end{array}, \quad p + q = 1$$

$$(k = 1, 2, \ldots, n).$$

Подсчитаем непосредственно дисперсию величины λ_k, учитывая, что центр распределения ее вероятностей равен p:

$$\sigma^2(\lambda_k) = \mathbf{M}(\lambda_k - p)^2 = (1-p)^2 p + (0-p)^2 q = q^2 p + p^2 q = pq.$$

Отсюда находим среднее квадратическое отклонение

$$\sigma(\lambda_k) = \sqrt{pq} \qquad (k = 1, 2, \ldots, n).$$

С помощью формулы (3.21) теперь получаем:

$$\sigma(\omega_n) = \frac{\sqrt{pq}}{\sqrt{n}}. \tag{3.22}$$

Отсюда с помощью формулы (3.18) можно найти также

$$\sigma(\mu_n) = \sigma(n\omega_n) = n\sigma(\omega_n) = \sqrt{npq}.$$

2. *Среднее квадратическое отклонение случайной величины* ξ, *равномерно распределенной в интервале* $(\alpha_1; \alpha_2)$. Центр распределения мы уже нашли ранее:

$$a = \mathbf{M}\xi = \frac{\alpha_1 + \alpha_2}{2}.$$

Подсчитаем теперь дисперсию непосредственно:

$$\sigma^2(\xi) = \mathbf{M}\left(\xi - \frac{\alpha_1 + \alpha_2}{2}\right)^2 = \int_{\alpha_1}^{\alpha_2} \left(x - \frac{\alpha_1 + \alpha_2}{2}\right)^2 \frac{1}{\alpha_2 - \alpha_1}\, dx = \frac{(\alpha_2 - \alpha_1)^2}{12}.$$

Отсюда находим, что среднее квадратическое отклонение

$$\sigma(\xi) = \frac{\alpha_2 - \alpha_1}{2\sqrt{3}}$$

пропорционально длине интервала $(\alpha_1; \alpha_2)$ (и составляет около трети этой длины).

3. *Дисперсия нормального распределения*. Для случайной величины ξ_0 с простейшим нормальным распределением вероятностей (2.22) центр $\mathbf{M}\xi_0 = 0$, и поэтому дисперсия равна

$$\mathbf{D}\xi_0 = \mathbf{M}\xi_0^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}}\, dx = 1\ ^*). \tag{3.23}$$

---

*) Интеграл этот удобно вычислять методом интегрирования по частям:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x\, x e^{-\frac{x^2}{2}}\, dx = \frac{1}{\sqrt{2\pi}} \left[-x e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\, dx = 1.$$

Для случайной величины $\xi = a + \sigma\xi_0$ с общим нормальным распределением (2.26) дисперсия равна

$$\mathbf{D}\xi = \mathbf{D}(a + \sigma\xi_0) = \sigma^2 \mathbf{D}\xi_0 = \sigma^2.$$

Отсюда следует, что $\sigma(\xi) = \sigma$. Вместе с ранее установленной формулой $\mathbf{M}\xi = a$ это позволяет полностью выяснить смысл параметров a и σ общего нормального распределения (2.26): a есть центр распределения, σ^2 — дисперсия. Рис. 7 (стр. 48) наглядно иллюстрирует роль параметра σ как характеристики рассеяния случайной величины с общим нормальным распределением вероятностей (при $a = 0$).

Минимальное свойство центра

Средний квадрат отклонения случайной величины ξ от центра распределения $a = \mathbf{M}\xi$ меньше, чем средний квадрат ее отклонения от любого другого числа:

$$\mathbf{M}(\xi - a)^2 < \mathbf{M}(\xi - C)^2 \qquad (C \neq a).$$

Доказательство. Так как $\mathbf{M}(\xi - a) = 0$, то

$$\begin{aligned}\mathbf{M}(\xi - C)^2 &= \mathbf{M}[(\xi - a) + (a - C)]^2 = \\ &= \mathbf{M}(\xi - a)^2 + 2(a - C)\,\mathbf{M}(\xi - a) + (a - C)^2 = \\ &= \mathbf{M}(\xi - a)^2 + (a - C)^2 \end{aligned} \qquad (3.24)$$

и, значит,

$$\mathbf{M}(\xi - C)^2 \geqslant \mathbf{M}(\xi - a)^2,$$

причем знак равенства достигается только при $(a - C)^2 = 0$, то есть при $C = a$.

Доказанное свойство указывает на важную связь между центром распределения и дисперсией: центр распределения минимизирует средний квадрат отклонения $\mathbf{M}(\xi - C)^2$, причем минимум этого среднего квадрата отклонения равен как раз дисперсии $\sigma^2(\xi)$.

Полученная нами формула (3.24) часто применяется для вычисления дисперсий*). В частности, при $C = 0$ эта формула дает:

$$\sigma^2(\xi) = \mathbf{M}(\xi - a)^2 = \mathbf{M}\xi^2 - a^2. \qquad (3.25)$$

*) Стоит обратить внимание на аналогию формулы (3.24) с соответствующей теоремой о моментах инерции.

Для примера вычислим дисперсию распределения Пуассона (2.5). Здесь проще сначала подсчитать $\mathsf{M}\xi^2$:

$$\mathsf{M}\xi^2 = \mathsf{M}\xi(\xi-1) + \mathsf{M}\xi = \sum_{m=0}^{\infty} m(m-1)\frac{a^m}{m!}e^{-a} + a =$$

$$= a^2 e^{-a} \sum_{m=2}^{\infty} \frac{a^{m-2}}{(m-2)!} + a = a^2 + a,$$

а затем найти дисперсию

$$\sigma^2(\xi) = \mathsf{M}\xi^2 - a^2 = (a^2 + a) - a^2 = a.$$

Полезно заметить, что в распределении Пуассона и центр распределения и дисперсия совпадают со значением параметра a.

Понятие о моментах распределения

Рассмотренные выше две основные характеристики распределения — центр распределения $\mathsf{M}\xi = a$ и дисперсия $\mathsf{M}(\xi - a)^2 = \sigma^2$ — представляют собой частные случаи моментов распределения, введенных известным русским математиком П. Л. Чебышевым для исследования законов распределения вероятностей.

Начальным моментом порядка k называется математическое ожидание k-й степени случайной величины, то есть $\mathsf{M}\xi^k$.

Центральным моментом порядка k называется математическое ожидание k-й степени отклонения случайной величины от ее центра распределения, то есть $\mathsf{M}(\xi - a)^k$.

Между начальными и центральными моментами существуют простые зависимости, легко устанавливаемые с помощью бинома Ньютона. Например:

$$\mathsf{M}(\xi - a)^2 = \mathsf{M}\xi^2 - 2a\mathsf{M}\xi + a^2 = \mathsf{M}\xi^2 - a^2;$$

$$\mathsf{M}(\xi - a)^3 = \mathsf{M}\xi^3 - 3a\mathsf{M}\xi^2 + 3a^2\mathsf{M}\xi - a^3 = \mathsf{M}\xi^3 - 3a\mathsf{M}\xi^2 + 2a^3$$

и т. д. Первая из этих формул совпадает с формулой (3.25).

Выше было отмечено, что моменты первого и второго порядков $\mathsf{M}\xi$ и $\mathsf{M}(\xi - a)^2$ характеризуют центр расположения и рассеяние случайной величины ξ. Центральный момент третьего порядка $\mathsf{M}(\xi - a)^3$ применяется для характеристики *асимметрии распределения*. Если кривая распределения симметрична относительно прямой $x = a$, то центральный момент

третьего порядка (как и вообще все центральные моменты нечетных порядков) будет равен нулю *). Поэтому если центральный момент третьего порядка отличен от нуля, то распределение не может быть симметричным. Величину асимметрии характеризуют обычно безразмерным *коэффициентом асимметрии* $C_s = \frac{M(\xi - a)^3}{\sigma^3(\xi)}$. Знак коэффициента асимметрии указывает на правостороннюю или левостороннюю асимметрию (рис. 18).

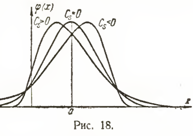


Рис. 18.

Моменты более высоких порядков в элементарных задачах теории вероятностей и в простейших ее приложениях не встречаются.

Упражнения

1. Подсчитать математическое ожидание произведения характеристических случайных величин λ_1 и λ_2, введенных для бесповторной выборки в примере на стр. 40, и убедиться в том, что в этом случае теорема умножения математических ожиданий неприменима.

Ответ.

$$M(\lambda_1\lambda_2) = \frac{25}{100} \cdot \frac{24}{99}; \qquad M\lambda_1 M\lambda_2 = \left(\frac{25}{100}\right)^2.$$

2. Доказать теорему умножения математических ожиданий для непрерывных независимых случайных величин с помощью формулы (3.13).

Указание. Воспользоваться условием независимости $\varphi(x; y) = \psi_1(x)\psi_2(y)$ и свести двойной интеграл $\iint (xy)\psi_1(x)\psi_2(y)\,dx\,dy$ к двум простым интегралам.

3. Найти центр распределения и среднее квадратическое отклонение для распределения (2.2) числа очков.

Ответ.

$$M\xi = \frac{1+2+3+4+5+6}{6} = 3{,}5; \quad \sigma = \sqrt{\frac{35}{12}} = 1{,}71.$$

*) Это следует из того, что при указанном условии плотность распределения $\psi(y)$ случайной величины $\eta = \xi - a$ (отклонения) будет четной функцией и, значит, все произведения $y^{2k+1}\psi(y)$ будут нечетными функциями.

4. То же для суммы чисел очков, выпадающих на двух игральных костях.

Ответ.

$$M\xi = 7; \quad \sigma = \sqrt{\frac{70}{12}} = 2{,}42.$$

Указание. Воспользоваться теоремами сложения математических ожиданий и дисперсий.

5. Найти математическое ожидание числа белых шаров в испытании по схеме упражнения 2 к главе II.

Ответ.

$$M\xi = \frac{5}{6}.$$

Указание. Представить ξ в виде суммы характеристических случайных величин, связанных с каждым выниманием шара.

6. Найти центр и дисперсию распределения (2.4) числа произведенных выстрелов при стрельбе по схеме примера 3 (стр. 31). Рассмотреть числовой пример при $p = \frac{1}{10}$ и дать толкование математического ожидания.

Ответ.

$$M\xi = \sum_{n=1}^{\infty} n(1-p)^{n-1} p = \frac{1}{p};$$

$$\sigma^2 = M\xi^2 - (M\xi)^2 = \sum_{n=1}^{\infty} n^2 (1-p)^{n-1} p - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

При $p = \frac{1}{10}$; $M\xi = 10$; если вероятность попадания при каждом выстреле равна $\frac{1}{10}$, то в среднем надо произвести 10 выстрелов для первого попадания.

Указание. Воспользоваться степенными рядами для

$$\frac{1}{(1-q)^2} \text{ и } \frac{1}{(1-q)^3}.$$

7. Найти центр и дисперсию распределения Пирсона (2.27).

Ответ.

$$M\xi = \int_0^{\infty} x \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\, dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}} = \frac{\alpha}{\beta};$$

$$\sigma^2 = M\xi^2 - (M\xi)^2 = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{\beta^{\alpha+2}} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}.$$

Указание. Воспользоваться интегрированием по частям или основным свойством гамма-функции.

8. Доказать теорему сложения центральных моментов третьего порядка для независимых случайных величин ξ, η:

$$\mathbf{M}\,[(\xi+\eta)-(a+b)]^3=\mathbf{M}\,(\xi-a)^3+\mathbf{M}\,(\eta-b)^3.$$

9. Найти коэффициент асимметрии биномиального распределения частоты μ_n.

Ответ.

$$\frac{\mathbf{M}\,(\mu_n-np)^3}{\sigma^3\,(\mu_n)}=\frac{npq\,(q-p)}{(npq)^{3/2}}=\frac{q-p}{\sqrt{npq}}.$$

Указание. Вычислить сначала третий центральный момент для характеристической величины λ:

$$\mathbf{M}\,(\lambda-p)^3=(1-p)^3\,p+(0-p)^3\,q=pq\,(q-p),$$

затем воспользоваться теоремой сложения для центральных моментов третьего порядка.

10. Доказать, что коэффициент асимметрии распределения Пирсона (2.27) вдвое больше так называемого коэффициента вариации

$$C_v=\frac{\sigma\,(\xi)}{\mathbf{M}\,\xi}=\frac{1}{\sqrt{\alpha}}.$$

Указание. При вычислении центрального момента $\mathbf{M}\,(\xi-a)^3$ выразить его через начальные моменты.

11. Вычислить центральный момент четвертого порядка для общего нормального распределения вероятностей (2.26).

Ответ.

$$\mathbf{M}\,(\xi-a)^4=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}(x-a)^4\,e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx=3\sigma^4.$$

Указание. Заменить $x-a=t\sigma$ и проинтегрировать по частям.

ГЛАВА IV

ЗАКОН БОЛЬШИХ ЧИСЕЛ

Для каждой случайиой величииы иельзя предвидеть, какое оиа примет зиачение в итоге испытаиия. Но поведеиие суммы большого числа случайиых величии почти утрачивает случайный характер и становится закоиомериым; здесь иеобходимое прокладывает себе дорогу сквозь миожество случайиостей. Теоремы, в которых выясияются условия того, что совокупиое действие миогих случайиых причии приводит к результату, почти ие зависящему от случая, получили общее иазваиие закоиа больших чисел.

§ 13. О случайных событиях с очень малыми вероятностями

Напомним, что вероятность p случайного события есть число, объективно характеризующее возможность его появления при данном комплексе условий. Относительная частота этого события есть случайная величина ω_n с распределением вероятностей, центром которого служит как раз вероятность p (§ 11, пример 3). В то время как значение вероятности не может быть найдено непосредственно из опыта, каждое n-кратное повторение испытания дает определенное, опытное значение относительной частоты. Мы уже указывали в начале нашего курса, что значение относительной частоты ω_n при достаточно большом количестве испытаний n оказывается, как правило, весьма близким к вероятности p. Этот общий принцип связывает теорию с практикой, но он оказывается слишком неопределенным для количественных оценок (мы знаем уже, что относительная частота близка к вероятности, но не знаем пока, насколько именно). Поэтому в основу дальнейших выводов мы положим другой, более узкий, но зато и более определенный принцип, относящийся к событиям с очеиь малыми вероятностями.

Если событие имеет очень малую вероятность, то оно происходит крайне редко. Например, если событие имеет вероятность 0,000001, то оно происходит приблизительно один раз на миллион испытаний *).

Опыт убеждает нас в том, что при малом количестве испытаний случайное событие с такой малой вероятностью, как правило, не происходит совсем; поэтому возможностью его появления мы пренебрегаем. Например, вряд ли кто-нибудь, обладая одним билетом в лотерее, где на 1 000 000 билетов приходится один выигрыш, станет всерьез рассчитывать на этот выигрыш (хотя один из миллиона обладателей таких билетов обязательно выиграет!). Ну, а если бы в лотерее было 500 000 билетов? или 10 000? Возникает вопрос, насколько мала должна быть вероятность случайного события, чтобы можно было пренебречь возможностью его появления в единичном испытании. Ответ на этот вопрос не может быть дан в теории вероятностей, он относится к ее практическим приложениям и зависит от существа решаемой проблемы. Поясним это следующим примером сравнения двух случайных событий.

а) Пусть при автоматической обработке некоторой детали вероятность получения нестандартного размера равна 0,01, причем нестандартная деталь будет браковаться при сборке. Если деталь недорогая, то вполне допустимо не производить сплошной контроль всех деталей перед сборкой, т. е. пренебречь вероятностью 0,01 нестандартности детали.

б) Пусть при изготовлении парашютов вероятность получения нераскрывающегося вовремя парашюта равна 0,01. Ясно, что в этом случае пренебречь вероятностью 0,01 недопустимо, так как это привело бы к гибели примерно каждого сотого парашютиста. В этом случае следовало бы организовать контроль всех изготовленных парашютов.

В каждой области приложения теории вероятностей назначается определенная граница «очень малых» вероятностей. В отношении этой границы принимается так называемый принцип практической невозможности

---

*) Разумеется, это отнюдь не означает, что оно происходит именно при миллионном испытании; оно может произойти и при одном из первых испытаний.

маловероятных событий: *считается, что случайное событие, имеющее вероятность менее назначенной границы, не произойдет при единичном испытании.*

Этот принцип играет основную роль в применениях теории вероятностей к практике; он позволит нам выяснить практическое значение рассматриваемых далее теорем.

Иногда этот принцип формулируют несколько иначе и называют «принципом практической уверенности»: если вероятность события A меньше назначенной границы α, то имеется практическая уверенность в том, что событие A не произойдет (при единичном испытании). При этом вероятность противоположного события $\overline{A}$ будет больше, чем $1-\alpha$, то есть будет так же близка к единице, как вероятность события A близка к нулю. Поэтому принцип практической уверенности формулируют также и следующим образом: если вероятность события больше, чем $1-\alpha$, то имеется практическая уверенность в том, что оно произойдет (при единичном испытании).

Сделаем еще следующее важное замечание о практическом приложении принятого выше принципа. Допустим, что, руководствуясь некоторой гипотезой, мы нашли, что вероятность события A очень мала (меньше назначенной границы). Но, произведя испытание, мы обнаружили, что событие A произошло. Тогда будет разумным подвергнуть сомнению нашу исходную гипотезу и искать неслучайную причину появления события A. Особенно ярко это выражено в следующем старинном анекдоте, приведенном Бертраном в его «Исчислении вероятностей» (1889 г.). «Однажды в Неаполе аббат Галлиани увидел человека из Базиликаты, который, встряхивая три игральные кости в чашке, держал пари, что выбросит три шестерки; и действительно, он немедленно получил три шестерки. Вы скажете, такая удача возможна *). Однако человеку из Базиликаты это удалось во второй раз **), и пари повторилось. Он клал кости назад в чашку три, четыре, пять раз, и каждый раз выбрасывал три шестерки. „Черт побери, — вскричал аббат, — кости налиты свинцом“. Так оно и было».

*) Вероятность этого при условии, что кость правильна, составляет:

$$\frac{1}{6^3}=\frac{1}{216}.$$

**) Вероятность этого при условии, что кость правильна, составляет уже

$$\frac{1}{(216)^2}=\frac{1}{46\,656}.$$

§ 14. Теорема Я. Бернулли и устойчивость относительных частот

Пусть при некотором испытании случайное событие A имеет определенную вероятность p. Пусть, далее, указанное испытание повторяется n раз *). Как мы уже знаем, относительная частота случайного события будет случайной величиной ω_n, у которой центр распределения совпадает с вероятностью p. Что же касается среднего квадратического отклонения величины ω_n, то оно будет уменьшаться с увеличением количества испытаний n, как показывает формула (3.22):

$$\sigma(\omega_n) = \frac{\sqrt{pq}}{\sqrt{n}}.$$

Отсюда следует, что с увеличением количества испытаний значения относительной частоты случайного события будут рассеиваться все меньше и меньше, то есть будут все теснее группироваться около вероятности этого события.

Уточнением этого положения является опубликованная в 1713 г. замечательная

Теорема Я. Бернулли. *Если в последовательности независимых испытаний вероятность p случайного события остается неизменной, то вероятность того, что отклонение относительной частоты ω_n от p превзойдет заданное число $\varepsilon > 0$, стремится к нулю при неограниченном увеличении числа испытаний n:*

$$\lim_{n \to \infty} \mathrm{P}\{|\omega_n - p| > \varepsilon\} = 0. \tag{4.1}$$

Стремление к нулю вероятности неравенства

$$|\omega_n - p| > \varepsilon$$

означает, что при достаточно большом n эта вероятность станет меньше назначенной границы очень малых вероятностей (см. § 13). При этом мы будем иметь практическую уверенность в том, что указанное неравенство не будет выполняться, и, следовательно, в том, что будет выполняться противоположное неравенство:

$$|\omega_n - p| \leqslant \varepsilon. \tag{4.2}$$

*) См. сноску к стр. 36.

Другими словами, теорему Я. Бернулли можно формулировать так:

При достаточно большом количестве испытаний достигается практическая уверенность в том, что отклонение относительной частоты случайного события от его вероятности не превзойдет по абсолютной величине наперед заданного числа ε (как бы мало оно ни было!).

Теорема Я. Бернулли является весьма частным случаем теоремы Чебышева, которую мы докажем далее, в § 15. Заметим, что, как бы велико ни было n, мы не можем категорически утверждать, что будет иметь место неравенство (4.2), а можем иметь только практическую уверенность, в смысле § 13, в выполнении этого неравенства. Чтобы подчеркнуть отличие установленного положения от обычного понятия предела, иногда вводят специальное понятие «*предела по вероятности*». Например, теорему Бернулли записывают в символическом виде:

$$\omega_n \xrightarrow[n \to \infty]{\text{вер}} p \tag{4.3}$$

(читается так: ω_n стремится по вероятности к числу p при $n \longrightarrow \infty$).

Если при n-кратном повторении испытания случайное событие A фактически появилось m раз, то отношение $\frac{m}{n}$ будет частным, опытным значением относительной частоты ω_n. При достаточно большом n можно быть практически уверенным в выполнении приближенного равенства

$$\frac{m}{n} \approx p \tag{4.4}$$

с какой угодно наперед заданной точностью. На практике это проявляется в том, что значения $\frac{m}{n}$ относительной частоты ω_n обладают устойчивостью, о которой мы говорили в начале книги.

Соотношение (4.4) может служить для приближенного вычисления неизвестной вероятности случайного события по опытным данным. Например, в статистике народонаселения XIX века было установлено, что относительная частота рождения мальчиков устойчива и составляет 0,512. Отсюда можно заключить,

что рождение мальчика имеет определенную вероятность, близкую к 0,512.

При конкретных значениях n точность приближенного равенства (4.4) нуждается в оценке; мы дадим эту оценку в следующей главе.

§ 15. Теорема Чебышева

Для независимых случайных величин общий закон больших чисел выражается теоремой П. Л. Чебышева, доказанной им в 1867 г.

Рассмотрим последовательность попарно независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ с какими угодно распределениями вероятностей. Допустим, что все эти случайные величины имеют определенные математические ожидания и дисперсии:

$$\mathbf{M}\xi_k = a_k; \qquad \mathbf{M}(\xi_k - a_k)^2 = \sigma_k^2 \qquad (k = 1, 2, \ldots). \quad (4.5)$$

Образуем среднюю арифметическую из первых n случайных величин:

$$\overline{\xi}_n = \frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n}. \quad (4.6)$$

Математическим ожиданием величины $\overline{\xi}_n$ служит среднее арифметическое из математических ожиданий величин $\xi_1, \xi_2, \ldots, \xi_n$:

$$\mathbf{M}\overline{\xi}_n = \frac{a_1 + a_2 + \ldots + a_n}{n} = \overline{a}_n. \quad (4.7)$$

Дисперсия же средней арифметической $\overline{\xi}_n$ не равна средней арифметической из дисперсий, а меньше ее в n раз:

$$\sigma^2(\overline{\xi}_n) = \frac{1}{n^2}(\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2) = \frac{1}{n}\,\frac{\sum \sigma_k^2}{n}. \quad (4.8)$$

Если все дисперсии σ_k^2 ограничены одним и тем же числом:

$$\sigma_k^2 \leqslant H \qquad (k = 1, 2, \ldots),$$

то дисперсия средней арифметической стремится к нулю при $n \to \infty$, так как

$$\sigma^2(\overline{\xi}_n) \leqslant \frac{H}{n}.$$

Отсюда следует, что с увеличением n значения средней арифметической $\bar{\xi}_n$ будут рассеиваться все меньше и меньше, то есть будут все теснее группироваться около своего центра распределения *).

Для уточнения этого положения оценим возможные отклонения случайной величины $\bar{\xi}_n$ от ее центра с помощью универсального неравенства Чебышева.

Неравенство Чебышева

Неравенство Чебышева дает оценку вероятности того, что отклонение любой случайной величины ξ от центра ее распределения $a = \mathbf{M}\xi$ превзойдет заданное положительное число ε:

$$\mathbf{P}\{|\xi - a| > \varepsilon\} < \frac{\sigma^2(\xi)}{\varepsilon^2}. \tag{4.9}$$

Оказывается, что эта вероятность, вообще говоря, тем меньше, чем меньше дисперсия $\sigma^2(\xi)$.

Проведем доказательство неравенства Чебышева для непрерывных случайных величин. По основной формуле (2.20) имеем:

$$\mathbf{P}\{|\xi - a| > \varepsilon\} = \int\limits_{|x-a|>\varepsilon} \varphi(x)\,dx,$$

где интеграл в правой части распространяется на интервалы $(-\infty;\ a-\varepsilon)$ и $(a+\varepsilon;\ +\infty)$, в которых $|x-a| > \varepsilon$. Так как в этих интервалах имеет место неравенство

$$1 < \frac{(x-a)^2}{\varepsilon^2},$$

а значит, и неравенство

$$\varphi(x) \leqslant \frac{1}{\varepsilon^2}(x-a)^2\varphi(x),$$

то

$$\int\limits_{|x-a|>\varepsilon} \varphi(x)\,dx \leqslant \frac{1}{\varepsilon^2}\int\limits_{|x-a|>\varepsilon} (x-a)^2\varphi(x)\,dx.$$

---

*) Это можно толковать таким образом, что при образовании средней арифметической происходит частичное взаимное погашение случайных отклонений разных знаков.

Теперь для завершения вывода неравенства (4.9) достаточно заметить, что

$$\int\limits_{|x-a|>\varepsilon} (x-a)^2 \varphi(x)\,dx \leqslant \int\limits_{-\infty}^{\infty} (x-a)^2 \varphi(x)\,dx = \sigma^2(\xi).$$

Рекомендуем читателю в качестве упражнения провести аналогичное доказательство неравенства Чебышева для дискретных величин.

Теорема Чебышева

Применим теперь неравенство Чебышева (4.9) к случайной величине $\overline{\xi}_n$:

$$\mathbf{P}\{|\overline{\xi}_n - \overline{a}_n| > \varepsilon\} < \frac{\sigma^2(\overline{\xi}_n)}{\varepsilon^2} < \frac{H}{n\varepsilon^2}. \qquad (4.10)$$

Как бы мало ни было наперед заданное положительное число ε, всегда можно выбрать настолько большое n, чтобы правая часть неравенства (4.10) стала как угодно мала, в частности, меньше назначенной границы «очень малых» вероятностей. Тогда мы будем иметь практическую уверенность в том, что неравенство $|\overline{\xi}_n - \overline{a}_n| > \varepsilon$ не будет выполняться, и, следовательно, в том, что будет выполняться противоположное неравенство:

$$|\overline{\xi}_n - \overline{a}_n| \leqslant \varepsilon.$$

Таким образом, *при достаточно большом количестве независимых случайных величин достигается практическая уверенность в том, что отклонение их средней арифметической от ее центра распределения не превзойдет по абсолютной величине наперед заданного числа ε (как бы мало оно ни было!).*

В малом рассеянии средней $\overline{\xi}_n$ около центра ее распределения при больших значениях n и заключается закон больших чисел. Точное математическое выражение этого закона дает

Теорема Чебышева. *Если последовательность $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ состоит из попарно независимых случайных величин с ограниченными дисперсиями, то для средней арифметической $\overline{\xi}_n$ из первых n величин $\xi_1, \xi_2, \ldots, \xi_n$*

при любом положительном ε *имеет место соотношение*

$$\lim_{n\to\infty} \mathbf{P}\{|\bar{\xi}_n - \mathbf{M}\bar{\xi}_n| > \varepsilon\} = 0. \tag{4.11}$$

Формула (4.11) получается из неравенства (4.10) предельным переходом, так как правая часть неравенства (4.10) стремится к нулю при $n \to \infty$, а левая часть, будучи вероятностью, не может быть отрицательной.

Частный случай теоремы Чебышева

Если все случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ имеют одинаковые центры распределения $\mathbf{M}\xi_k = a$ $(k = 1, 2, \ldots)$, то центр распределения средней арифметической $\bar{\xi}_n$ также совпадает с a:

$$\mathbf{M}\bar{\xi}_n = \frac{1}{n}(\mathbf{M}\xi_1 + \mathbf{M}\xi_2 + \ldots + \mathbf{M}\xi_n) = a,$$

и поэтому формула (4.11) принимает вид

$$\lim_{n\to\infty} \mathbf{P}\{|\bar{\xi}_n - a| > \varepsilon\} = 0. \tag{4.12}$$

Формула (4.12) допускает символическую запись

$$\bar{\xi}_n \xrightarrow[n\to\infty]{\text{вер}} a, \tag{4.13}$$

вполне аналогичную записи (4.3) *).

Доказательство теоремы Я. Бернулли

Если в качестве величин ξ_k взять характеристические величины λ_k, введенные в § 7, то их средняя арифметическая будет равна относительной частоте случайного события:

$$\bar{\lambda}_n = \frac{\lambda_1 + \lambda_2 + \ldots + \lambda_n}{n} = \dot{\omega}_n.$$

Так как при этом $\mathbf{M}\lambda_k = p$; $\sigma^2(\lambda_k) = pq < 1$, то формула (4.12) принимает вид (4.1), что и доказывает теорему Я. Бернулли.

*) В 1928 г. чл.-корр. АН СССР А. Я. Хинчин доказал, что для одинаково распределенных независимых случайных величин формула (4.12) имеет место при одном только предположении о конечности их центра распределения a без всяких ограничений на дисперсию σ^2 (которая может быть и бесконечной).

§ 16. Устойчивость выборочных средних и метод моментов

Рассмотрим сначала статистическую задачу о средних значениях. Пусть из совокупности N элементов, различающихся некоторым количественным признаком x, отбираются случайным образом n элементов. Можно ли считать, что средние арифметические значения признака во всей совокупности и в отобранной совокупности будут близки между собой?

Теорема Чебышева дает решение этого вопроса при условии, что выборка является повторной, то есть производится по схеме «возвращенного шара», описанной на стр. 36. Свяжем с каждым отбираемым элементом случайную величину ξ_k, равную возможному значению признака у k-го элемента выборки. При условии повторности выборки отбор каждого элемента производится из одной и той же исходной совокупности, и поэтому все случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ будут независимы и будут иметь одинаковые распределения вероятностей типа (3.2). Как было показано в § 10, центр распределения всех этих величин совпадает со средним арифметическим значением признака в исходной («генеральной») совокупности, то есть с так называемой «генеральной средней» a:

$$\mathbf{M}\xi_k = a \quad (k = 1, \ 2, \ldots, \ n).$$

Поэтому для арифметической средней $\overline{\xi}_n = \frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n}$ будет справедлива теорема Чебышева в форме (4.12) — (4.13). Это значит, что среднее возможное значение признака в выборке стремится по вероятности к генеральной средней при неограниченном увеличении объема выборки.

Сделаем отсюда практические выводы. Опытным значением каждой случайной величины ξ_k является то значение, которое мы фактически обнаруживаем у k-го элемента выборки; опытным значением случайной величины $\overline{\xi}_n$ служит *выборочная средняя* $\overline{x}$ — среднее арифметическое значение признака у отобранных элементов. Поэтому соотношение (4.13) может быть истолковано так: *при достаточно большом объеме* (n) *случайной повторной выборки можно быть практически уверенным в том, что выборочная средняя будет сколь угодно мало отличаться от генеральной средней*, то есть в том, что будет

иметь место приближенное равенство

$$\bar{x} \approx a. \tag{4.14}$$

Отсюда следует также, что *выборочные средние обладают устойчивостью,* то есть в двух случайных повторных выборках достаточно большого объема выборочные средние должны быть приближенно равны. Этот вывод хорошо согласуется с опытом.

Подчеркнем, что степень близости выборочной средней к генеральной зависит только от объема выборки n и не зависит от отношения объема выборки к объему генеральной совокупности *). Заметим еще, что если объем генеральной совокупности очень велик по сравнению с объемом выборки, то повторность выборки оказывается несущественной и наши выводы можно применять по отношению к бесповторным случайным выборкам, что особенно важно на практике. Дело в том, что генеральная средняя, вообще говоря, неизвестна и о ее значении судят по величине выборочной средней. Насколько это может быть важно, видно из следующего примера. Допустим, что для некоторых расчетов надо знать среднюю продолжительность службы большой партии электроламп. Чтобы точно узнать среднюю продолжительность службы для всей партии, надо подвергнуть испытанию все лампы; но тогда у нас не останется ни одной лампы, для которой имела бы значение полученная средняя. Практически такие средние, как среднюю продолжительность службы ламп, находят выборочным методом, то есть с помощью случайной выборки. Конечно, приближенное определение генеральной средней по выборочной средней нуждается в оценке точности; эту оценку мы дадим в главе VI.

О методе моментов

Приближенному равенству (4.14) можно дать и другое толкование.

Пусть ξ — некоторая случайная величина с конечным центром распределения $M\xi = a$. Будем производить независимые

*) Например, 1%-ная выборка из совокупности в 1 000 000 элементов дает более точные сведения о генеральной средней, чем 2%-ная выборка из совокупности в 1000 элементов (при той же величине σ).

испытания, в результате которых величина ξ примет определенные значения $x_1, x_2, \ldots, x_n$. Эти опытные значения величины ξ можно рассматривать и как значения *разных* случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ с тем же распределением вероятностей, что и у величины ξ, причем величины $\xi_1, \xi_2, \ldots, \xi_n$ можно считать независимыми в силу независимости испытаний. При таком толковании среднее арифметическое из опытных данных

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

можно рассматривать как опытное значение случайной величины $\overline{\xi_n}$, для которой справедлива теорема Чебышева в форме (4.12). Поэтому для достаточно больших n можно ожидать выполнения с достаточной точностью приближенного равенства

$$\overline{x} \approx \mathbf{M}\xi = a. \tag{4.15}$$

Отсюда следует, что *приближенным значением математического ожидания случайной величины является среднее арифметическое из ее значений, полученных опытным путем.*

Это положение позволяет приближенно находить не только центр распределения, но и другие моменты распределения, которые определяются тоже как математические ожидания некоторых величин. Например, для дисперсии мы получаем такую приближенную формулу:

$$\sigma^2(\xi) = \mathbf{M}(\xi - a)^2 \approx \frac{\sum (x_k - a)^2}{n}, \tag{4.16}$$

где сумма берется по всем опытным данным $x_1, x_2, \ldots, x_n$.

Действительно, значение $\frac{\sum (x_k - a)^2}{n}$ можно рассматривать как частное значение средней арифметической из n одинаково распределенных независимых случайных величин $(\xi_k - a)^2$ с математическим ожиданием

$$\mathbf{M}(\xi_k - a)^2 = \mathbf{M}(\xi - a)^2 = \sigma^2(\xi) \qquad (k = 1, 2, \ldots, n).$$

Поэтому

$$\mathbf{P}\left\{\left|\frac{\sum (\xi_k - a)^2}{n} - \sigma^2(\xi)\right| > \varepsilon\right\} \to 0 \quad \text{при} \quad n \to \infty.$$

В формулу (4.16) входит значение центра распределения a. Обычно это значение неизвестно и естественно попытаться заменить его приближенным значением $\bar{x}$. Оказывается, однако, что получающаяся при этом формула

$$\sigma^2(\xi) \approx \frac{\sum (x_k - \bar{x})^2}{n} \tag{4.17}$$

уже *не будет верна* (в том смысле, в каком верны приближенные формулы (4.15) и (4.16)). Дело в том, что хотя значение $\frac{\sum (x_k - \bar{x})^2}{n}$ можно и здесь рассматривать как частное значение средней арифметической из n случайных величин $(\xi_k - \bar{\xi}_n)^2$, но математическим ожиданием этих величин уже не будет служить $\sigma^2(\xi)$. Действительно, непосредственный подсчет математического ожидания дает:

$$\mathbf{M}(\xi_1 - \bar{\xi}_n)^2 = \mathbf{M}\left[(\xi_1 - a) - \left(\frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n} - a\right)\right]^2 =$$
$$= \mathbf{M}\left[(\xi_1 - a)\left(1 - \frac{1}{n}\right) - \left(\frac{\xi_2 - a}{n}\right) - \ldots - \left(\frac{\xi_n - a}{n}\right)\right]^2 =$$
$$= \sigma^2(\xi_1)\left(1 - \frac{1}{n}\right)^2 + \frac{\sigma^2(\xi_2)}{n^2} + \ldots + \frac{\sigma^2(\xi_n)}{n^2} = \frac{n-1}{n}\sigma^2(\xi);$$

и аналогично для любого k: $\mathbf{M}(\xi_k - \bar{\xi}_n)^2 = \frac{n-1}{n}\sigma^2(\xi)$ *). В силу линейности математического ожидания отсюда следует, что случайные величины $\frac{n}{n-1}(\xi_k - \bar{\xi}_n)^2$ будут иметь своим математическим ожиданием как раз $\sigma^2(\xi)$; и хотя эти величины не будут независимы, закон больших чисел будет для них справедлив **); поэтому значения их средней арифметической

$$\frac{\sum \frac{n}{n-1}(\xi_k - \bar{\xi}_n)^2}{n} = \frac{\sum (\xi_k - \bar{\xi}_n)^2}{n-1}$$

*) Причина этого заключается в линейной зависимости между величинами $\xi_1, \xi_2, \ldots, \xi_n$ и их средней $\bar{\xi}_n = \frac{1}{n}(\xi_1 + \xi_2 + \ldots + \xi_n)$.

**) Доказательство этого опускаем.

будут мало отличаться от их центра $\sigma^2(\xi)$ при достаточно большом числе n.

Таким образом, формулу (4.17) можно исправить введением в правую часть множителя $\frac{n}{n-1}$. Исправленная приближенная формула для вычисления дисперсии по опытным данным имеет вид

$$\sigma^2(\xi) \approx \frac{\sum (x_k - \bar{x})^2}{n-1}. \qquad (4.18)$$

Величина, стоящая в правой части этого соотношения, называется *выборочной дисперсией* и обозначается через s_n^2. Полезно заметить, что при больших значениях n число $n-1$ относительно мало отличается от n и поэтому формулы (4.17) и (4.18) дают практически одинаковые результаты. Но при малых значениях n различие между этими формулами весьма заметно.

Конечно, все приведенные выше приближенные формулы при конкретном n нуждаются в оценке погрешности. Некоторые оценки мы укажем в главе VI.

Возможность приближенного нахождения моментов распределения по опытным данным позволяет решать задачу нахождения параметров распределения, если тип распределения известен.

Поясним эту задачу на примерах рассмотренных выше распределений.

1. Для общего нормального распределения с плотностью

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

параметры a и σ^2 являются центром распределения и дисперсией (см. § 12). Они могут быть найдены из опытных данных непосредственно по формулам (4.15) и (4.18).

2. Для равномерного распределения вероятностей неизвестными параметрами могут быть начало и конец интервала возможных значений $(\alpha_1; \alpha_2)$.

Как мы знаем из § 12, моменты распределения выражаются через эти параметры по формулам:

$$a = \mathrm{M}\xi = \frac{\alpha_1 + \alpha_2}{2}; \quad \sigma = \sigma(\xi) = \frac{\alpha_2 - \alpha_1}{2\sqrt{3}}.$$

Отсюда находим:

$$\alpha_1 = a - \sigma\sqrt{3}; \quad \alpha_2 = a + \sigma\sqrt{3},$$

где a и σ^2 находятся по формулам (4.15) и (4.18).

3. Для распределения Пирсона (2.27) параметры α и β связаны с центром и дисперсией формулами (см. упражнение 7 к главе III):

$$a = \mathbf{M}\xi = \frac{\alpha}{\beta}; \quad \sigma^2 = \sigma^2(\xi) = \frac{\alpha}{\beta^2}.$$

Отсюда находим:

$$\alpha = \frac{a^2}{\sigma^2}; \quad \beta = \frac{a}{\sigma^2},$$

где a и σ^2 находятся по формулам (4.15) и (4.18)

Если известно, что плотность распределения зависит от l неизвестных параметров $\alpha_1, \alpha_2, \ldots, \alpha_l$, то, выражая через них первые l моментов распределения, мы получаем l уравнений, из которых (вообще говоря) можем найти все l параметров. Сами же моменты распределения могут быть найдены по опытным данным, как указывалось выше. Следует заметить, однако, что чем выше порядок момента распределения, тем больше нужно опытных данных для его сколько-нибудь точного вычисления. Поэтому на практике часто ограничиваются распределениями только с двумя неизвестными параметрами, которые находят с помощью первых двух моментов распределения.

Упражнения

1. На телефонной станции производились наблюдения за числом ξ неправильных соединений в минуту. Наблюдения в течение часа дали следующие результаты:

3	1	3	1	4	2
2	4	0	3	0	2
2	0	2	1	4	3
3	1	4	2	2	1
1	2	1	0	3	4
1	3	2	7	2	0
0	1	3	3	1	2
4	2	0	2	3	1
2	5	1	1	0	1
1	2	2	1	1	5

Найти центр и дисперсию распределения и проверить выполнение основного условия для распределения Пуассона $\mathbf{M}\xi = \sigma^2 = a$. Найти

соответствующее распределение Пуассона. Сравнить таблицу распределения опытных данных с соответствующей таблицей распределения Пуассона.

Ответ. Среднее число неправильных соединений в минуту равно $\overline{x}=2$; условие для распределения Пуассона $s_n^2 \approx 2{,}1 = \overline{x}$. Распределение Пуассона для задачи:

$$\mathrm{P}\{\xi=m\}=\frac{2^m}{m!}e^{-2} \quad (a=2).$$

Сравнение опытных данных с расчетными:

Число неправильных соединений	Частота	Относительная частота	Вероятность по распределению Пуассона
0	8	0,1333	0,1353
1	17	0,2833	0,2707
2	16	0,2667	0,2707
3	10	0,1667	0,1804
4	6	0,1000	0,0902
5	2	0,0333	0,0361
$\geqslant 6$	1	0,0167	0,0166
	60	1,0000	1,0000

2. Измерения 100 обработанных деталей дали следующие отклонения от номинального размера:

−2	2	1	2	−1	−2	3	1	−1	0
0	−1	3	1	2	−3	1	0	1	1
0	1	−1	1	0	2	2	1	0	−1
1	1	4	−1	1	1	−1	0	2	−2
2	0	−2	0	0	−1	1	4	−2	1
−3	0	0	1	4	0	−2	2	1	2
−1	1	0	−1	0	3	1	−2	3	−1
1	2	2	0	−2	1	0	−1	0	3
3	−2	−1	−2	1	0	0	−3	1	0
2	1	0	3	−1	2	1	0	−1	0

Найти центр и дисперсию распределения и построить соответствующее нормальное распределение вероятностей. Сравнить таблицу функции распределения по опытным данным (таблицу накопленных относительных частот) с соответствующей таблицей функции нормального распределения (см. стр. 49).

Ответ. Среднее отклонение $\overline{x}=0{,}4$; выборочная дисперсия $s_n^2=2{,}57$. Соответствующее нормальное распределение имеет $a=0{,}4$, $\sigma=1{,}6$ и, значит, плотность

$$\varphi(x)=\frac{1}{1{,}6\sqrt{2\pi}}e^{-\frac{(x-0{,}4)^2}{2\cdot 2{,}57}}.$$

Сравнение опытных данных с расчетными:

Отклонения x	Частота m	Накопленная относительная частота	$F(x+0{,}5)$
−3	3	0,03	0,0349
−2	10	0,13	0,1175
−1	15	0,28	0,2869
0	24	0,52	0,5249
1	25	0,77	0,7541
2	13	0,90	0,9053
3	7	0,97	0,9736
4	3	1,00	0,9948
	100		

Здесь следует считать каждое значение x средним для соответствующего интервала; например, частоту $m=10$ относить к интервалу $-2{,}5<x<-1{,}5$; это обстоятельство мы учитываем при сравнении с функцией нормального распределения, беря ее в виде

$$F(x+0{,}5)=\frac{1}{2}+\frac{1}{2}\Phi\left(\frac{x+0{,}5-a}{\sigma}\right),$$

где $\Phi(t)$ — интеграл вероятностей.

3. Найти распределение Пирсона (2.27) для следующей таблицы опытных данных (в приведенных единицах):

Значения x	0	1	2	3	4	5	6	
Частота m	1	33	41	18	5	1	1	100

Проверить выполнение основного условия для распределения Пирсона $C_s=2C_v$ (см. упражнение 7 к главе III). Сравнить таблицу накопленных частот по опытным данным с таблицей соответствующей функции распределения Пирсона.

Ответ. $\overline{x}=2{,}0$; $s_n^2=1{,}0$. Полагаем для распределения Пирсона (2.27): $a=\frac{\alpha}{\beta}=2$; $\sigma^2=\frac{\alpha}{\beta^2}=1$, откуда $\alpha=4$; $\beta=2$. Соответствующее распределение Пирсона имеет плотность $\varphi(x)=\frac{16}{3!}x^3e^{-2x}$. Условие для распределения Пирсона:

$$C_v=\frac{\sigma}{a}=0{,}5; \qquad C_s=1{,}09\approx 2C_v.$$

Сравнение опытных данных с расчетными:

x	m	Накопленная относительная частота	$F(x+0{,}5)$
0	1	0,01	0,0190
1	33	0,34	0,3527
2	41	0,75	0,7350
3	18	0,93	0,9183
4	5	0,98	0,9788
5	1	0,99	0,9951
6	1	1,00	0,9990

Здесь $F(x+0{,}5)=\frac{16}{3!}\int\limits_0^{x+0{,}5}x^3e^{-2x}dx.$

ГЛАВА V

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ И ОЦЕНКИ СРЕДНИХ

Для оценки относительных частот и некоторых других средних величин решающее значение имеет то обстоятельство, что распределение вероятностей этих величин стремится к нормальному распределению. Для доказательства соответствующих предельных теорем известный русский математик А. М. Ляпунов разработал весьма мощный метод характеристических функций, который позволил ему доказать так называемую центральную предельную теорему (см. § 19). Прежде чем приводить предельные теоремы, сообщим здесь необходимые сведения о характеристических функциях.

§ 17. Понятие о характеристических функциях

Характеристической функцией $f(u)$ случайной величины ξ называется математическое ожидание величины $e^{iu\xi}$:

$$f(u) = \mathrm{M}e^{iu\xi}, \tag{5.1}$$

где u — действительный параметр.

Для дискретной случайной величины

$$f(u) = \sum e^{iux_k} p_k, \tag{5.2}$$

где p_k есть вероятность значения x_k и сумма берется по всем значениям x_k величины ξ. Для непрерывной случайной величины

$$f(u) = \int_{-\infty}^{\infty} e^{iux} \varphi(x)\, dx, \tag{5.3}$$

где $\varphi(x)$ — плотность распределения величины ξ. Интеграл (5.3) всегда сходится, и притом абсолютно, так как $|e^{iux}\varphi(x)| = \varphi(x)$, а интеграл от $\varphi(x)$ сходится; при этом

$$|f(u)| \leqslant \int_{-\infty}^{\infty} \varphi(x)\, dx = 1.$$

Основные свойства характеристических функций

1. Характеристическая функция однозначно определяет распределение вероятностей случайной величины *). Другими словами, если две случайные величины имеют одинаковые характеристические функции, то они имеют также и одинаковые распределения вероятностей.

2. Если характеристическая функция $f(u)$ *непрерывной* случайной величины ξ является пределом последовательности характеристических функций $f_n(u)$ каких угодно случайных величин ξ_n $(n = 1, 2, 3, \ldots)$, то функция распределения $F(x) = \mathbf{P}\{\xi < x\}$ является пределом последовательности функций распределения $F_n(x) = \mathbf{P}\{\xi_n < x\}$; таким образом, из $\lim\limits_{n\to\infty} f_n(u) = f(u)$ следует, что $\lim\limits_{n\to\infty} F_n(x) = F(x)$ для всех x **).

Это свойство важно тем, что во многих случаях предельный переход в последовательности характеристических функций осуществляется проще, чем в последовательности функций распределения. Поэтому доказательство предельных теорем с помощью характеристических функций оказывается в этих случаях более коротким и простым.

Указанные выше свойства мы примем без доказательства.

3. Характеристическая функция суммы независимых случайных величин равна произведению характеристических функций слагаемых.

Докажем это свойство для двух независимых случайных величин ξ и η с характеристическими функциями $f_\xi(u)$ и $f_\eta(u)$.

Так как случайные величины $e^{iu\xi}$ и $e^{iu\eta}$ также будут независимыми, то для вычисления характеристической функции $f_{\xi+\eta}(u)$ для суммы $\xi + \eta$ можно применить теорему умножения математических ожиданий

$$f_{\xi+\eta}(u) = \mathbf{M}e^{iu(\xi+\eta)} = \mathbf{M}e^{iu\xi}e^{iu\eta} = \mathbf{M}e^{iu\xi}\mathbf{M}e^{iu\eta},$$

то есть

$$f_{\xi+\eta}(u) = f_\xi(u) f_\eta(u). \tag{5.4}$$

Таким образом, нахождение характеристической функции суммы независимых случайных величин проще, чем нахождение соответствующего распределения вероятностей (которое сводится к свертке плотностей распределения слагаемых, см. § 9).

4. При линейном преобразовании случайной величины, то есть при переходе от случайной величины ξ к случайной величине $\eta = A + B\xi$, характеристическая функция преобразуется по формуле

$$f_\eta(u) = e^{iAu} f_\xi(Bu). \tag{5.5}$$

Эта формула проверяется непосредственно:

$$\mathbf{M}e^{iu\eta} = \mathbf{M}e^{iu(A+B\xi)} = e^{iAu}\mathbf{M}e^{i(Bu)\xi}.$$

*) Можно даже указать общую формулу выражения функции распределения через характеристическую функцию; см. Б. В. Гнеденко, Курс теории вероятностей, ГИТТЛ, 1954, глава 7. Там же можно найти доказательства формулируемых нами свойств. Для читателей, знакомых с интегралом Фурье, заметим, что интеграл (5.3) есть преобразование Фурье от плотности $\varphi(x)$.

**) Более общие теоремы этого рода см. Б. В. Гнеденко, Курс теории вероятностей.

Примеры характеристических функций:

1) Для случайной величины λ $\begin{array}{|c|c|} 1 & 0 \\ \hline p & q \end{array}$ характеристическая функция легко подсчитывается по формуле (5.2):

$$f_\lambda(u) = e^{iu1}p + e^{iu0}q = pe^{iu} + q.$$

2) Частота случайного события есть сумма

$$\mu_n = \lambda_1 + \lambda_2 + \ldots + \lambda_n,$$

где все λ_k независимы и имеют одно и то же распределение λ_k $\begin{array}{|c|c|} 1 & 0 \\ \hline p & q \end{array}$ $(k = 1, 2, \ldots, n)$ (см. § 7). Характеристическую функцию частоты μ_n подсчитаем по свойству 3:

$$f_{\mu_n}(u) = f_{\lambda_1}(u) f_{\lambda_2}(u) \ldots f_{\lambda_n}(u) = (pe^{iu} + q)^n. \tag{5.6}$$

3) Относительная частота случайного события есть $\omega_n = \frac{\mu_n}{n}$; характеристическую функцию ее подсчитаем по формулам (5.5) и (5.6):

$$f_{\omega_n}(u) = f_{\mu_n}\left(\frac{u}{n}\right) = \left(pe^{i\frac{u}{n}} + q\right)^n.$$

4) Случайная величина, равномерно распределенная в интервале $(-a;\ a)$, имеет плотность

$$\varphi(x) = \begin{cases} \frac{1}{2a} & (-a < x < a), \\ 0 & (|x| > a). \end{cases}$$

Ее характеристическую функцию подсчитаем по формуле (5.3)

$$f(u) = \int_{-a}^{a} e^{iux} \frac{1}{2a}\, dx = \frac{e^{iua} - e^{-iua}}{2aiu} = \frac{\sin au}{au}. \tag{5.7}$$

5) Случайная величина ξ_0 с простейшим нормальным распределением (2.22) имеет характеристическую функцию *)

$$f_0(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} e^{-\frac{x^2}{2}}\, dx = e^{-\frac{u^2}{2}}. \tag{5.8}$$

*) Вычисление интеграла (5.8) мы опускаем.

Случайная величина ξ с общим нормальным распределением (2.26) связана со случайной величиной ξ_0 линейной зависимостью $\xi = a + \sigma\xi_0$. Ее характеристическую функцию подсчитаем по формулам (5.8) и (5.5):

$$f(u) = e^{iau} f_0(\sigma u) = e^{iau} e^{-\frac{\sigma^2 u^2}{2}}. \tag{5.9}$$

Отсюда следует, в частности, что если взаимно независимые случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ имеют общие нормальные распределения с центрами $a_1, a_2, \ldots, a_n$ и дисперсиями $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, то их сумма имеет нормальное распределение с центром $a = a_1 + a_2 + \ldots + a_n$ и дисперсией $\sigma^2 = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2$; действительно, из $f_k(u) = e^{ia_k a} e^{-\frac{\sigma_k^2 u^2}{2}}$ $(k = 1, 2, \ldots, n)$ следует $f(u) = f_1(u) f_2(u) \ldots f_n(u) = e^{iau} e^{-\frac{\sigma^2 u^2}{2}}$.

Связь между характеристической функцией и моментами распределения

Так как характеристическая функция $f(u)$ однозначно определяет распределение вероятностей случайной величины ξ, то все моменты распределения могут быть выражены через характеристическую функцию. Для получения этих выражений мы продифференцируем равенство

$$f(u) = \mathrm{M} e^{iu\xi}$$

формально по u (под знаком математического ожидания, то есть под знаком суммы или интеграла). Мы получим последовательно

$$f'(u) = \mathrm{M} i\xi e^{iu\xi},$$
$$f''(u) = \mathrm{M} (i\xi)^2 e^{iu\xi},$$
$$\cdots\cdots\cdots$$
$$f^{(k)}(u) = \mathrm{M} (i\xi)^k e^{iu\xi}.$$

Можно доказать, что такое дифференцирование законно, если случайная величина ξ имеет моменты до k-го порядка включительно. Полагая в полученных формулах $u = 0$, находим связь между производными характеристической функции в нуле и начальными моментами:

$$f(0) = \mathrm{M}1 = 1,$$
$$f'(0) = i\mathrm{M}\xi,$$
$$f''(0) = -\mathrm{M}\xi^2,$$
$$\cdots\cdots\cdots$$
$$f^{(k)}(0) = i^k \mathrm{M}\xi^k.$$

В приложениях встречаются также производные от логарифма характеристической функции: $\psi(u) = \ln f(u)$.

Число $i^k\psi^{(k)}(0)$ называется *семиинвариантом* k-го порядка случайной величины ξ. Легко проверить, что

$$i\psi'(0) = -\mathbf{M}\xi; \quad i^2\psi''(0) = \sigma^2(\xi).$$

Семиинварианты играют большую роль при рассмотрении сумм независимых случайных величин, так как при сложении независимых случайных величин их семиинварианты тоже складываются.

§ 18. Предельная теорема Муавра — Лапласа; оценка относительных частот

В настоящем параграфе мы рассмотрим предельное распределение относительной частоты случайного события при неограниченном увеличении числа испытаний. Как известно из § 7, распределение относительной частоты ω_n для повторной выборки является биномиальным, то есть

$$\mathbf{P}\left\{\omega_n = \frac{m}{n}\right\} = C_n^m p^m q^{n-m} \qquad (m = 0, 1, 2, \ldots, n), \quad (5.10)$$

где n — число (независимых) испытаний, p — вероятность рассматриваемого события в каждом испытании. Если число испытаний велико, то расчет вероятностей по формуле (5.10) становится весьма затруднительным. Характер затруднений станет еще более ясным, если учесть, что в практических приложениях нас интересует не вероятность отдельного равенства $\omega_n = \frac{m}{n}$, а вероятность неравенства $|\omega_n - p| < \varepsilon$, оценивающего отклонение относительной частоты ω_n от ее вероятностного предела (см. § 14). Вероятность указанного неравенства равна сумме $\sum C_n^m p^m q^{n-m}$, распространенной на те значения m, для которых $\left|\frac{m}{n} - p\right| < \varepsilon$, то есть на значения m, заключенные между $np - n\varepsilon$ и $np + n\varepsilon$. Возникающие здесь вычислительные трудности поясним примером. Пусть нас интересует вероятность того, что при 10 000 испытаний отклонение относительной частоты события от его вероятности $p = 0,2$ не превзойдет $\varepsilon = 0,01$. Здесь $n = 10\,000$, $np = 2000$, $q = 0,8$, и для точного подсчета интересующей нас вероятности нам придется вычислить сумму более двухсот слагаемых вида $\frac{10\,000!}{m!\,(10\,000 - m)!}(0,2)^m (0,8)^{10\,000 - m}$ для m от $np - n\varepsilon = 1900$ до $np + n\varepsilon = 2100$. Непосредственное вычисление такой суммы со сколько-нибудь удовлетворительной точностью требует

огромной затраты труда. Поэтому уже давно возникла задача приближенного подсчета вероятностей с помощью замены точного биномиального распределения величины ω_n некоторым предельным непрерывным распределением*). Указанная задача была успешно решена Муавром (в 1730 г.) для частного случая $p=q=\frac{1}{2}$, а затем Лапласом (в 1783 г.) для общего случая любого p, $0<p<1$. Оказывается, что для биномиального распределения существует предельное (при $n\longrightarrow\infty$) распределение и это предельное распределение является нормальным.

Для удобства формулировки соответствующей теоремы предварительно пронормируем относительную частоту ω_n.

Нормированием случайной величины ξ *называется линейное преобразование ее в новую величину*

$$\xi_0=\frac{\xi-\mathbf{M}\xi}{\sigma(\xi)}.$$

Это преобразование сводится к переносу начала отсчета в центр распределения ($\mathbf{M}\xi$) и выбору в качестве единицы масштаба среднего квадратического отклонения $\sigma(\xi)$.

Нормированную относительную частоту обозначим через τ_n:

$$\tau_n=\frac{\omega_n-\mathbf{M}\omega_n}{\sigma(\omega_n)}=\frac{\omega_n-p}{\sqrt{\frac{pq}{n}}}. \tag{5.11}$$

Теорема Муавра — Лапласа

При неограниченном увеличении числа испытаний предельным распределением вероятностей нормированной относительной частоты случайного события является простейшее нормальное распределение, то есть

$$\lim_{n\to\infty}\mathbf{P}\{|\tau_n|<t\}=\Phi(t), \tag{5.12}$$

где $\Phi(t)$ *есть интеграл вероятностей* (2.23).

Эта теорема является частным случаем более общей теоремы, которую мы докажем далее, в § 19. Здесь же мы укажем лишь характер применений теоремы Муавра — Лапласа.

*) Вычисление вероятностей неравенств в непрерывном распределении сводится к вычислению интеграла, что обычно значительно проще, чем вычисление сумм для дискретного распределения.

Применение теоремы Муавра — Лапласа к оценке относительных частот

Теорема Муавра — Лапласа позволяет оценить вероятность неравенства $|\omega_n - p| < \varepsilon$ при достаточно больших n (и при значениях p, не слишком близких к 0 или 1). Возьмем настолько большое n, чтобы с удовлетворяющей нас точностью можно было считать, что

$$\mathbf{P}\{|\tau_n| < t\} \approx \Phi(t). \tag{5.13}$$

Тогда из равносильности неравенств

$$|\omega_n - p| < \varepsilon \quad \text{и} \quad |\tau_n| = \left|\frac{\omega_n - p}{\sqrt{\frac{pq}{n}}}\right| < \frac{\varepsilon\sqrt{n}}{\sqrt{pq}}$$

вытекает, что вероятность интересующего нас неравенства $|\omega_n - p| < \varepsilon$ приближенно равна интегралу вероятностей

$$\mathbf{P}\{|\omega_n - p| < \varepsilon\} \approx \Phi(t), \qquad \text{где} \qquad t = \frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \tag{5.14}$$

Примеры. Понятие о доверительных оценках. Для числового примера, приведенного в начале настоящего параграфа, $p = 0,2$; $q = 0,8$; $n = 10\,000$ и $\varepsilon = 0,01$. Вероятность интересующего нас неравенства $|\omega_n - 0,2| < 0,01$ вычисляем приближенно по формуле (5.14):

$$t = \frac{0,01\sqrt{10\,000}}{\sqrt{0,2\cdot 0,8}} = 2,5;$$

$$\mathbf{P}\{|\omega_n - 0,2| < 0,01\} \approx \Phi(2,5) = 0,988\,^{*)}.$$

Если считать вероятность $P = 0,988$ достаточно близкой к единице, то можно быть практически уверенным в том, что $|\omega_n - 0,2| < 0,01$, то есть что при десяти тысячах испытаний по схеме повторной выборки относительная частота случайного события отклонится от его вероятности $p = 0,2$ менее чем на 0,01. Вероятность 0,988 в этом случае

---

*) Здесь приближенная формула верна до третьего знака, что объясняется довольно большим значением n (10 000). Точные оценки погрешности формулы (5.14) можно найти в специальных статьях С. Н. Бернштейна, В. Феллера и др. Заметим, что если число n имеет порядок лишь нескольких сотен (причем np и nq все же значительно больше 1), то вместо формулы (5.14) рекомендуется применять не-

называется «*надежностью*» *оценки*

$$|\omega_n - 0{,}2| < 0{,}01,$$

а сама эта оценка называется *доверительной оценкой относительной частоты* ω_n с надежностью 0,988.

На практике надежность оценки задается заранее в соответствии с назначенным пределом «очень малых вероятностей» (см. § 13). Например, если мы решили пренебречь возможностью появления события с вероятностью 0,001, то назначаем надежность $P = 1 - 0{,}001 = 0{,}999$.

По заданной надежности P находим соответствующее значение $t = t(P)$ из уравнения

$$\Phi(t) = P$$

с помощью таблиц интеграла вероятностей; например, для $P = 0{,}999$ находим $t = 3{,}29$. Тогда доверительная оценка с заданной надежностью P принимает вид

$$|\omega_n - p| < \varepsilon, \quad \text{где} \quad \varepsilon = t(P)\sqrt{\frac{pq}{n}}. \qquad (5.15)$$

Неравенство (5.15) означает, что относительная частота ω_n с заданной надежностью P должна лежать в интервале

сколько более точную формулу

$$\mathbf{P}\left\{|\omega_n - p| \leqslant \frac{k}{n}\right\} \approx \Phi\left(\frac{k + \frac{1}{2}}{\sqrt{npq}}\right) \qquad (k \text{ — целое}).$$

Какова точность этой последней формулы, покажем на примерах.

$$1)\ p = \frac{1}{2};\quad n = 200;\quad k = 5;$$

$$\mathbf{P}\left\{\left|\omega_n - \frac{1}{2}\right| \leqslant \frac{5}{200}\right\} = 0{,}56325;$$

$$\frac{1}{\sqrt{npq}} = \frac{1}{\sqrt{50}} = 0{,}14142;\quad \Phi\left(\frac{5{,}5}{\sqrt{npq}}\right) = 0{,}56331.$$

2) Для несимметричного интервала точность меньше; например, при

$$p = 0{,}1;\quad n = 500;\quad k = 5;\quad \mathbf{P}\{0{,}1 \leqslant \omega_n \leqslant 0{,}11\} = 0{,}3176$$

$$\frac{1}{\sqrt{npq}} = \frac{1}{\sqrt{45}} = 0{,}1491;\quad \frac{1}{2}\Phi\left(\frac{5{,}5}{\sqrt{npq}}\right) - \frac{1}{2}\Phi\left(\frac{-0{,}5}{\sqrt{npq}}\right) = 0{,}3235.$$

Здесь расчет ведем по формуле типа (2.24).

$(p-\varepsilon,\ p+\varepsilon)$, где $\varepsilon=t(P)\sqrt{\frac{pq}{n}}$. Такой интервал называется «*доверительным интервалом*».

Для рассмотренного выше примера доверительная оценка относительной частоты с надежностью $P=0,999$ будет:

$$|\omega_n-0,2|<3,29\cdot\sqrt{\frac{0,2\cdot 0,8}{10\,000}}=0,0132.$$

Это значит, что с надежностью $P=0,999$ мы можем ожидать, что относительная частота ω_n будет лежать в доверительном интервале (0,1868; 0,2132). Для частоты $\mu_n=n\omega_n$ доверительный интервал будет соответственно в n раз больше: $1868<\mu_n<2132$ (при $n=10\,000$).

Практическая ценность доверительных интервалов заключается не только в возможности заранее предсказать границы частот (или границы относительных частот). Если испытания произведены в действительности и при этом оказалось, что фактическая частота вышла за доверительные границы, то это заставляет нас подвергнуть сомнению правильность вычисления вероятности интересующего нас случайного события. Такая постановка вопроса оказывается полезной, например, в деле регулирования массового производственного процесса. Поясним это следующим схематическим примером. Пусть автоматическая обработка некоторой детали отрегулирована так, что доля нестандартных деталей не превышает 1%. Для того чтобы проверить, не происходит ли увеличение этой доли, то есть не происходит ли нарушение установленного процесса, можно произвести выборочную проверку. При выборке *) в n деталей частота $\mu_n=n\omega_n$ нестандартных деталей должна лежать в доверительном интервале

$$n(p-\varepsilon)<\mu_n<n(p+\varepsilon),$$

то есть не должна превышать числа $np+n\varepsilon$, где $\varepsilon=t(P)\sqrt{\frac{pq}{n}}$, P — заданная надежность, p — предполагаемая вероятность получения нестандартной детали. Если,

*) Выборка должна быть случайной и повторной. Но если объем выборки очень мал по сравнению с объемом всей партии деталей, то и для бесповторной выборки указанные формулы дают достаточно хорошую точность.

например, $p=0{,}01\ (=1\%)$, $n=1000$ и $P=0{,}999$, то $np=10$; $t=3{,}29$, $n\varepsilon=3{,}29\sqrt{1000\cdot 0{,}01\cdot 0{,}99}=10{,}4$; $np+n\varepsilon=20{,}4$ и, значит, количество нестандартных деталей в выборке не должно превышать двадцати. Если же в произведенной выборке количество нестандартных деталей окажется больше двадцати, то надо заключить, что производственный процесс нарушен и доля нестандартных деталей превысила допуск в 1%.

§ 19. Доверительные оценки средних. Понятие о центральной предельной теореме Ляпунова

Нормальность предельного распределения в теореме Муавра — Лапласа связана не с какими-либо специфическими свойствами биномиального распределения, а лишь с тем обстоятельством, что относительная частота ω_n есть средняя арифметическая для независимых случайных величин $\lambda_1, \lambda_2, \ldots, \lambda_n$. Эта теорема допускает непосредственное обобщение на средние арифметические для любой последовательности независимых случайных величин с одинаковыми распределениями вероятностей (в предположении конечности их центра и дисперсии).

Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — последовательность взаимно независимых случайных величин, каждая из которых имеет одно и то же распределение вероятностей с центром $\mathbf{M}\xi_k=a$ и дисперсией $\mathbf{M}(\xi_k-a)^2=\sigma^2\ (k=1, 2, 3, \ldots)$. Образуем средние арифметические для первых n величин:

$$\bar{\xi}_n=\frac{\xi_1+\xi_2+\ldots+\xi_n}{n};$$

для удобства формулировки соответствующей теоремы введем еще нормированные средние (или нормированные суммы):

$$\tau_n=\frac{\bar{\xi}_n-\mathbf{M}\bar{\xi}_n}{\sigma(\bar{\xi}_n)}=\frac{(\xi_1+\xi_2+\ldots+\xi_n)-\mathbf{M}(\xi_1+\xi_2+\ldots+\xi_n)}{\sigma(\xi_1+\xi_2+\ldots+\xi_n)}. \quad (5.16)$$

В силу независимости случайных величин ξ_k имеем (см. главу III) $\mathbf{M}\bar{\xi}_n=a$; $\sigma(\bar{\xi}_n)=\frac{\sigma}{\sqrt{n}}$.

Поэтому

$$\tau_n = \frac{\bar{\xi}_n - a}{\frac{\sigma}{\sqrt{n}}} = \frac{(\xi_1 + \xi_2 + \ldots + \xi_n) - na}{\sigma\sqrt{n}}.$$

Теорема. *Предельным* (*при* $n \to \infty$) *распределением нормированных средних* (5.16) *является нормальное распределение, то есть*

$$\lim_{n \to \infty} \mathbf{P}\{|\tau_n| < t\} = \Phi(t), \qquad (5.17)$$

где $\Phi(t)$ — *интеграл вероятностей* (2.23).

Доказательство этой теоремы проведем методом характеристических функций, причем ограничимся непрерывными случайными величинами. Обозначим через $\varphi(x)$ плотность распределения для нормированной случайной величины $\xi_k^{(0)} = \frac{\xi_k - a}{\sigma}$ (эта плотность по условию одна и та же для всех величин ξ_k). При этом

$$\int \varphi(x)\,dx = 1,$$

$$\int x\varphi(x)\,dx = \mathbf{M}\xi_k^{(0)} = \frac{\mathbf{M}(\xi_k - a)}{\sigma} = 0,$$

$$\int x^2\varphi(x)\,dx = \mathbf{M}[\xi_k^{(0)}]^2 = \frac{\mathbf{M}(\xi_k - a)^2}{\sigma^2} = 1.$$

Обозначим, далее, через $f(u)$ характеристическую функцию, общую для всех нормированных величин $\xi_k^{(0)}$:

$$f(u) = \int e^{iux}\varphi(x)\,dx.$$

Рассмотрим теперь последовательность характеристических функций $f_n(u)$ для случайных величин τ_n. Выразим $f_n(u)$ через $f(u)$, для чего сначала выразим величину τ_n через $\xi_k^{(0)}$:

$$\tau_n = \frac{1}{\sqrt{n}} \frac{(\xi_1 - a) + (\xi_2 - a) + \ldots + (\xi_n - a)}{\sigma} =$$

$$= \frac{1}{\sqrt{n}}(\xi_1^{(0)} + \xi_2^{(0)} + \ldots + \xi_n^{(0)}).$$

Пользуясь свойствами 3 и 4 характеристических функций и независимостью случайных величин $\xi_k^{(0)}$, находим.

$$f_n(u) = \left[f\left(\frac{u}{\sqrt{n}}\right)\right]^n,$$

то есть

$$f_n(u) = \left[\int e^{i\frac{u}{\sqrt{n}}x} \varphi(x)\,dx \right]^n.$$

Разлагая, далее, подынтегральную функцию в ряд по степеням $\frac{1}{\sqrt{n}}$ и ограничиваясь членами порядка $\frac{1}{n}$, получаем:

$$\int e^{i\frac{u}{\sqrt{n}}x} \varphi(x)\,dx = \int \left(1 + i\frac{u}{\sqrt{n}}x - \frac{u^2}{2n}x^2 + \ldots\right) \varphi(x)\,dx =$$
$$= \int \varphi(x)\,dx + i\frac{u}{\sqrt{n}} \int x\varphi(x)\,dx - \frac{u^2}{2n} \int x^2\varphi(x)\,dx + \ldots,$$

то есть

$$\int e^{i\frac{u}{\sqrt{n}}x} \varphi(x)\,dx = 1 - \frac{u^2}{2n} + \frac{\alpha_n}{n},$$

где $\alpha_n \to 0$ при $n \to \infty$ *).

Теперь легко найти предел последовательности

$$f_n(u) = \left(1 - \frac{u^2}{2n} + \frac{\alpha_n}{n}\right)^n.$$

По известным правилам математического анализа находим:

$$\lim_{n \to \infty} f_n(u) = e^{-\frac{u^2}{2}}.$$

Таким образом, последовательность характеристических функций $f_n(u)$ сходится к характеристической функции $f_0(u) = e^{-\frac{u^2}{2}}$ простейшего нормального распределения (см. § 17). Отсюда следует, что последовательность функций распределения для нормированных средних τ_n сходится к функции распределения простейшего нормального закона, а это равносильно утверждению (5.17).

Доказанная теорема позволяет получить доверительные оценки средних, то есть найти точность (ε) для неравенства $|\bar{\xi}_n - a| < \varepsilon$ с заданной надежностью P. С этой целью мы заменяем неравенство $|\bar{\xi}_n - a| < \varepsilon$ равносильным, а

---

*) Если интегралы берутся по конечному интервалу, то стремление α_n к нулю при $n \to \infty$ очевидно; для несобственных интегралов это требует специальных оценок, которые мы опускаем.

значит, и равновероятным ему неравенством

$$\left|\frac{\xi_n - a}{\frac{\sigma}{\sqrt{n}}}\right| < \frac{\varepsilon\sqrt{n}}{\sigma}, \text{ то есть } |\tau_n| < t = \frac{\varepsilon\sqrt{n}}{\sigma}.$$

Так как по доказанному

$$\lim_{n\to\infty} \mathbf{P}\{|\tau_n| < t\} = \Phi(t),$$

то при достаточно большом n вероятность неравенства $|\tau_n| < t$ будет приближенно равна $\Phi(t)$.

Следовательно, вероятность интересующего нас неравенства $|\bar{\xi}_n - a| < \varepsilon$ будет также приближенно равна $\Phi(t)$, где $t = \frac{\varepsilon\sqrt{n}}{\sigma}$. Задаваясь теперь определенной вероятностью P, которую мы считаем достаточно близкой к 1, мы находим по таблицам интеграла вероятностей значение $t = t(P)$, удовлетворяющее уравнению $\Phi(t) = P$, и получаем *доверительную оценку средней* ξ_n:

$$|\bar{\xi}_n - a| < \varepsilon = t(P)\frac{\sigma}{\sqrt{n}} \quad \text{с надежностью } P. \quad (5.18)$$

Отклонения опытной средней от математического ожидания. Пусть нас интересует *одна* случайная величина ξ с центром a и дисперсией σ^2 и пусть для определения ее частных значений производится достаточно большое число n независимых испытаний. Каковы бы ни были полученные при этом опытные значения $x_1, x_2, \ldots, x_n$, мы можем с вероятностью P утверждать, что их средняя арифметическая

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

будет удовлетворять неравенству

$$|\bar{x} - a| < t(P)\frac{\sigma}{\sqrt{n}}, \quad (5.19)$$

то есть будет лежать в доверительном интервале $(a - \varepsilon;\ a + \varepsilon)$, где $\varepsilon = t(P)\frac{\sigma}{\sqrt{n}}$. Это утверждение вытекает из оценки (5.18),

если связать с каждым k-м испытанием случайную величину ξ_k с тем же распределением вероятностей, что и у ξ; при этом частным значением ξ_k будет x_k, а частным значением $\overline{\xi}_n$ будет $\overline{x}$ (независимость же случайных величин ξ_k следует из предположенной независимости испытаний).

Отклонения выборочной средней от генеральной. Если рассматривать возможные значения признака у каждого элемента выборки как случайную величину с распределением (3.2), то выборочная средняя $\overline{x}$ будет значением средней арифметической из указанных величин. Поэтому она будет удовлетворять неравенству (5.19), где a есть генеральная средняя (как центр распределения (3.2)), а

$$\sigma = \sqrt{(x_1 - a)^2 \frac{M_1}{N} + (x_2 - a)^2 \frac{M_2}{N} + \ldots + (x_\nu - a)^2 \frac{M_\nu}{N}}.$$

Другими словами, с вероятностью P можно ожидать, что выборочная средняя $\overline{x}$ отклонится от генеральной средней a не более чем на $t(P)\frac{\sigma}{\sqrt{n}}$.

Доверительная оценка средней может быть использована при контроле и регулировании производственного процесса, в котором нужно поддерживать значение некоторого параметра, например размера детали, в определенных жестких границах допуска; если при выборочной проверке средних какое-либо среднее значение выйдет за границы доверительного интервала $(a - \varepsilon;\ a + \varepsilon)$, где $\varepsilon = t(P)\frac{\sigma}{\sqrt{n}}$, то надо будет проверить, не нарушен ли установленный режим. Более точная разработка подобных принципов привела к созданию специальных методов так называемого «статистического контроля качества».

Понятие о центральной предельной теореме Ляпунова

Выше мы установили, что нормальное распределение вероятностей является предельным для нормированных средних или, что то же самое, для нормированных сумм одинаково распределенных слагаемых.

Центральная предельная теорема устанавливает общие условия, при которых предельным распределением нормированных сумм взаимно независимых случайных слагаемых будет нормальное распределение. Эта проблема в общей форме впервые была поставлена в исследованиях П. Л. Чебышева, но полученные им условия были довольно ограничительными. При весьма общих условиях центральная предельная теорема была доказана в 1900 г. А. М. Ляпуновым, в связи с чем эта теорема и носит его имя. Ляпунов доказал достаточность следующих двух условий:

а) все случайные слагаемые имеют конечные абсолютные центральные моменты третьего порядка

$$\mathbf{M}\,|\xi_k - a_k|^3 \qquad (a_k = \mathbf{M}\xi_k; \qquad k = 1,\ 2,\ \ldots);$$

б) отношение

$$\frac{\sum\limits_{k=1}^{n} \mathbf{M}\,|\xi_k - a_k|^3}{\left[\sum\limits_{k=1}^{n} \sigma^2(\xi_k)\right]^{\frac{3}{2}}} \to 0 \quad \text{при} \quad n \to \infty. \tag{5.20}$$

Заметим, что для одинаково распределенных слагаемых условие б) выполняется автоматически, так как при этом отношение (5.20) принимает вид

$$\frac{n\,\mathbf{M}\,|\xi - a|^3}{[n\sigma^2]^{\frac{3}{2}}} = \frac{1}{\sqrt{n}}\,\frac{\mathbf{M}\,|\xi - a|^3}{\sigma^3}.$$

Смысл условий Ляпунова состоит в «предельной пренебрежимости» отдельными слагаемыми при образовании суммы, в равномерно малом влиянии на сумму каждого отдельного слагаемого.

Это еще отчетливее видно в несколько более общих условиях Линдеберга, в которых требуется равномерная малость вероятностей больших отклонений $|\xi_k - a_k|$ по сравнению с дисперсией суммы $\sigma^2(\xi_1 + \xi_2 + \ldots + \xi_n)$. Грубо говоря, среди слагаемых не должно быть таких, возможные отклонения которых доминировали бы над возможными отклонениями всех остальных.

Центральная предельная теорема Ляпунова позволила объяснить широкое распространение нормального закона распре-

деления в природе и в технике тем, что рассеяние изучаемых величин вызывается очень большим количеством случайных причин, влияние каждой из которых ничтожно мало. С другой стороны, установление точных условий центральной предельной теоремы позволяет строго ограничить область применимости нормального закона распределения.

В заключение надо отметить, что важность закона больших чисел и центральной предельной теоремы для всей теории вероятностей и ее приложений вызывает многочисленные исследования по их уточнению и обобщению в разных направлениях. Ведущая роль в этих исследованиях со времен Чебышева и до наших дней принадлежит русским и советским математикам. В частности, следует отметить исследования в области зависимых случайных величин, начатые еще учеником П. Л. Чебышева — А. А. Марковым и продолженные затем советскими математиками С. Н. Бернштейном, Е. Е. Слуцким и др. Теорема Чебышева, например, оказывается справедливой и для последовательности зависимых случайных величин $\xi_1, \xi_2, \ldots$ с ограниченными дисперсиями, если только связь между ними достаточно быстро убывает с увеличением разности номеров *). Подобные же условия ослабления связи позволяют распространить на зависимые величины и центральную предельную теорему Ляпунова.

Упражнения

1. Доказать теорему Муавра—Лапласа непосредственно с помощью характеристической функции частоты μ_n (5.6).

Указание. С помощью формул (5.6), (5.5) и (5.11) построить характеристическую функцию $f_n(u)$ нормированной частоты $\tau_n = \frac{\mu_n - np}{\sqrt{npq}}$ и затем с помощью разложения в ряд экспонент

$$e^{i\frac{u}{\sqrt{n}}\sqrt{\frac{q}{p}}} \quad \text{и} \quad e^{-i\frac{u}{\sqrt{n}}\sqrt{\frac{p}{q}}}$$

преобразовать характеристическую функцию к виду

$$f_n(u) = \left(1 - \frac{u^2}{2n} - i\frac{u^3}{3!\,n\sqrt{n}}\,\frac{q-p}{\sqrt{pq}} + \ldots\right)^n.$$

---

*) Точное условие можно записать с помощью коэффициента корреляции (см. § 25); для справедливости теоремы Чебышева оказывается достаточным, чтобы

$$r(\xi_i; \xi_k) \to 0 \quad \text{при} \quad |i-k| \to \infty.$$

2. Оценить относительную частоту ω_n при $p=0{,}01$; $n=1000$ с надежностью $P=0{,}99$. Сформулировать вывод.

Ответ.

$$|\omega_n - p| < 2{,}576 \sqrt{\frac{0{,}0099}{1000}} = 0{,}0081,$$

то есть $0{,}0019 < \omega_n < 0{,}0181$. С надежностью 0,99 можно ожидать, что при тысячекратном повторении испытания интересующее нас событие произойдет от 2 до 18 раз $(2 \leqslant \mu_n = n\omega_n \leqslant 18)$.

3. Было произведено 12 000 бросаний монеты, при этом герб выпадал 6019 раз. Насколько хорошо согласуется это с предположением о том, что вероятность выпадения герба равна $\frac{1}{2}$?

Ответ. При указанном предположении вероятность наблюденного или большего отклонения равна

$$\mathrm{P}\left\{\left|\omega_n - \frac{1}{2}\right| \geqslant \frac{19}{12\,000}\right\} \approx 1 - \Phi\left(\frac{\frac{19}{12\,000}\sqrt{12\,000}}{\sqrt{0{,}5\cdot 0{,}5}}\right) = 0{,}738.$$

Эта вероятность не мала, так что произведенный опыт не дает основания сомневаться в высказанной гипотезе о вероятности выпадения герба.

4. Распределение некоторого признака в большой партии изделий дается таблицей:

Значения признака x	3,40	3,45	3,50	3,55	3,60	3,65	3,70	3,75	
Частота M	150	380	1320	1530	970	470	100	80	5000

Производится случайная выборка в 100 изделий. Оценить выборочную среднюю с надежностью $P=0{,}99$.

Ответ. Генеральная средняя $a=3{,}55$; $\sigma = 0{,}05\cdot\sqrt{1{,}844} = 0{,}068$. Оценка выборочной средней при $P=0{,}99$:

$$|\bar{x} - a| < 2{,}576\cdot\frac{0{,}068}{\sqrt{100}} = 0{,}0175,$$

то есть

$$3{,}5325 < \bar{x} < 3{,}5675.$$

5. Случайная выборка из некоторой партии изделий дала следующее распределение признака:

Значение признака x	3,40	3,45	3,50	3,55	3,60	3,65	3,70	3,75	
Частота m	3	5	12	28	28	14	8	2	100

Можно ли считать, что среднее значение признака в этой партии не отличается от среднего значения в партии предыдущего примера (если принять условие неизменности σ в разных партиях)?

Ответ. Здесь $\overline{x}=3{,}55+0{,}05\cdot\frac{57}{100}=3{,}55+0{,}0285$. Отклонение от генеральной средней из предыдущего примера составляет $\varepsilon=0{,}0285$. Вероятность таких или бо́льших отклонений равна

$$\mathbf{P}\{|\overline{x}-a|\geqslant 0{,}0285\}\approx 1-\Phi\left(\frac{0{,}0285\sqrt{100}}{0{,}068}\right)<0{,}00003.$$

Эта вероятность очень мала, поэтому нельзя считать, что выборка произведена из партии со средним значением $a=3{,}55$ (при условии $\sigma=0{,}068$).

6. Доказать, что распределение Пуассона можно рассматривать как предельное для биномиального распределения при $n\to\infty$ и $p\to 0$, если $np=a$ остается неизменным.

Указание. Заменив $p=\frac{a}{n}$, перейти в формуле

$$C_n^m p^m q^{n-m}=\frac{n(n-1)\dots(n-m+1)}{m!}\frac{a^m}{n^m}\left(1-\frac{a}{n}\right)^{n-m}$$

к пределу при $n\to\infty$.

7. При сложении большого числа n слагаемых, округленных до единиц одного и того же разряда 10^{-m}, принимается, что погрешность округления каждого слагаемого есть случайная величина (ξ) с равномерным распределением вероятностей в интервале $(-0{,}5\cdot 10^{-m};+0{,}5\cdot 10^{-m})$. Показать, что абсолютная погрешность суммы с вероятностью 0,997 не превысит $\sqrt{3n}\cdot 0{,}5\cdot 10^{-m}$.

Указание. Погрешность суммы рассматривать как сумму n одинаково распределенных независимых слагаемых — погрешностей округления. Считать, что при достаточно большом n погрешность суммы имеет распределение, близкое к нормальному с центром 0 и средним квадратическим отклонением

$$\sigma=\sqrt{n}\,\sigma(\xi)=\sqrt{n}\,\frac{10^{-m}}{2\sqrt{3}}$$

(см. пример 2 § 12, стр. 74).

ГЛАВА VI

ПРИМЕНЕНИЕ ТЕОРИИ ВЕРОЯТНОСТЕЙ К МАТЕМАТИЧЕСКОЙ ОБРАБОТКЕ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ

§ 20. Случайные ошибки измерения, их распределение

Ошибкой или погрешностью измерения называется разность $x - a$ *между результатом измерения* x *и истинным значением измеряемой величины* a.

Всякое измерение сопряжено с ошибками. Повторяя измерение одной и той же величины даже в одинаковых условиях, мы получаем обычно различные результаты. По этим результатам мы должны судить об истинном значении измеряемой величины. Ясно, что как непосредственные результаты измерения, так и любой результат их обработки дают не точное, а лишь приближенные значения величины a. Из всех таких приближений надо выбрать в каком-то смысле наилучшее. Далее, надо оценить точность полученного приближения, то есть установить границу, которую заведомо (с заданной вероятностью) не превзойдет отклонение истинного значения от найденного приближения.

Применимость методов теории вероятностей к решению указанных задач основана на том, что возможный результат измерения является случайной величиной с определенным распределением вероятностей. Установим тип этого распределения в случае *прямых измерений* (когда значения измеряемой величины считываются непосредственно со шкалы прибора).

Мы будем считать, что *результаты измерения не содержат систематических ошибок.* Систематическая ошибка вызывается постоянно действующей причиной, и величина этой ошибки либо постоянна во всех измерениях, либо ме-

няется по определенному известному закону. Поэтому систематические ошибки могут быть устранены путем выверки и настройки измерительного прибора или введением соответствующих поправок к результатам измерения.

После устранения систематических ошибок результаты измерения все еще будут содержать неустранимые, неизбежные ошибки, которые получили название *случайных ошибок измерения* *). Эти ошибки вызываются многочисленными трудно уловимыми причинами, каждая из которых приводит лишь к незначительному колебанию результатов измерения (например, при взвешивании на аналитических весах к таким причинам относятся незначительные колебания температуры и влажности воздуха, колебания стола, попадание соринок на взвешиваемый предмет и т. д.).

Каждая из указанных причин порождает свою, так называемую *элементарную ошибку измерения;* очевидно, что реально наблюдаемая случайная ошибка является суммой этих элементарных ошибок. Если считать, что количество элементарных ошибок очень велико, а роль каждой из них в образовании реальной случайной ошибки очень мала **), то в силу центральной предельной теоремы случайная ошибка измерения должна более или менее точно следовать нормальному закону распределения вероятностей. Решающим доводом в пользу нормального закона распределения случайных ошибок является практическое подтверждение его анализом многочисленных опытов и наблюдений. Этот анализ показывает, что наблюдающееся распределение случайных ошибок измерения очень хорошо согласуется с нормальным законом, то есть что относительные частоты случайных ошибок определенной величины достаточно близки к вероятностям таких ошибок, рассчитанным по нормальному закону распределения.

В силу указанных причин в теории ошибок принимают в качестве основного постулата, что *при прямых измерениях случайная ошибка* τ *следует нормальному закону распре-*

---

*) Иногда в результате нарушения установленных условий измерения или при неправильной записи показаний прибора появляются еще и грубые ошибки (промахи). Соответствующие результаты измерения необходимо сразу же отбрасывать и не учитывать при дальнейшей обработке.

**) Это — так называемая «гипотеза элементарных ошибок».

деления вероятностей *). При этом, учитывая обычно наблюдающуюся симметрию положительных и отрицательных случайных ошибок, принимают еще, что *центр распределения случайных ошибок равен нулю*. Таким образом, плотность распределения вероятностей случайной ошибки τ равна

$$\varphi(t)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{t^2}{2\sigma^2}} \quad (-\infty<t<+\infty). \qquad (6.1)$$

Параметр $\sigma=\sqrt{\mathbf{M}\,\tau^2}$ называется *средней квадратической ошибкой измерения* или *стандартом*. Он характеризует точность измерений (или точность прибора).

Зная закон распределения случайных ошибок, легко найти закон распределения возможных результатов измерения, так как возможный результат измерения ξ и случайная ошибка τ связаны простой зависимостью:

$$\xi=a+\tau; \qquad (6.2)$$

при этом из $\mathbf{M}\tau=0$ следует $\mathbf{M}\,\xi=a$; это — так называемое условие несмещенности, которое практически связано с отсутствием систематических ошибок. Таким образом, возможный результат измерения ξ при указанных выше условиях следует нормальному закону распределения вероятностей с центром a и дисперсией σ^2.

§ 21. Решение двух основных задач теории ошибок. Оценка истинного значения измеряемой величины и оценка точности прибора в случае прямых равноточных измерений

Пусть $x_1, x_2, \ldots, x_n$ — результаты n прямых независимых измерений некоторой постоянной величины a. Мы предположим, что возможные результаты всех измерений

*) Надо заметить, что случайные ошибки, как и результаты измерения, всегда выражаются в некоторых целых единицах, связанных с шагом шкалы измерительного прибора; но в теории удобнее считать случайную ошибку *непрерывной* случайной величиной, что значительно упрощает все расчеты. Далее, в теории удобнее считать, что случайные ошибки распределены на всей оси, хотя иногда это условие даже противоречит физическому смыслу задачи (например, вес тела не может быть отрицательным). Практически условие неограниченности τ не сказывается на выводах, так как вероятность выхода τ за определенные границы очень мала.

$\xi_1, \xi_2, \ldots, \xi_n$ подчиняются нормальному закону распределения с одним и тем же центром

$$\mathbf{M}\xi_k = a \qquad (k = 1,\ 2,\ \ldots,\ n) \tag{6.3}$$

(условие несмещенности) и одной и той же дисперсией

$$\mathbf{M}(\xi_k - a)^2 = \sigma^2 \qquad (k = 1,\ 2,\ \ldots,\ n) \tag{6.4}$$

(условие равноточности измерений).

Как мы уже знаем из главы IV, § 16 (стр. 91), в качестве приближенного значения a целесообразно принять среднее арифметическое значение из результатов измерений:

$$a \approx \overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}. \tag{6.5}$$

Нашей первой задачей является оценка точности приближенного равенства (6.5).

Учитывая, что все случайные величины ξ_k независимы и имеют нормальное распределение вероятностей с центром a и дисперсией σ^2, мы находим (см. стр. 101 и 73), что средняя случайная величина

$$\overline{\xi} = \frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n}$$

также имеет нормальное распределение с центром a, но с дисперсией $\frac{\sigma^2}{n}$. Поэтому вероятность того, что средняя ошибка $\overline{\xi} - a$ не превзойдет по абсолютной величине некоторого положительного числа ε, равна

$$\mathbf{P}\{|\overline{\xi} - a| < \varepsilon\} = \Phi(t), \tag{6.6}$$

где $t = \frac{\varepsilon}{\sigma(\overline{\xi})} = \frac{\varepsilon\sqrt{n}}{\sigma}$, а $\Phi(t)$ есть интеграл вероятностей (2.23).

Обычно вероятность оценки (P) задается заранее и выбирается достаточно близкой к 1 (например, $P = 0{,}999$). Далее, из уравнения $\Phi(t) = P$ находится соответствующее значение t (по таблицам интеграла вероятностей; например, при $P = 0{,}999$ находим $t = 3{,}291$). Это приводит к оценке

$|\bar{\xi}-a|<\varepsilon$, где $\varepsilon=t\frac{\sigma}{\sqrt{n}}$. Подставляя в эту оценку вместо случайных величин ξ_k их опытные значения x_k и, значит, вместо средней $\bar{\xi}$ ее опытное значение $\bar{x}$, мы получим так называемую *классическую оценку*:

$$|\bar{x}-a|<t\frac{\sigma}{\sqrt{n}}$$

или

$$\bar{x}-t\frac{\sigma}{\sqrt{n}}<a<\bar{x}+t\frac{\sigma}{\sqrt{n}}. \qquad (6.7)$$

Оценку (6.7) надо понимать следующим образом: вероятность того, что интервал $(\bar{x}-\varepsilon;\ \bar{x}+\varepsilon)$, где $\varepsilon=t\frac{\sigma}{\sqrt{n}}$, будет заключать истинное значение a измеряемой величины, равна заранее заданному числу $P=\Phi(t)$. Число P называется *надежностью оценки* (6.7).

Классическая оценка (6.7) имеет тот существенный недостаток, что при этом предполагается известной дисперсия σ^2. Если же эту дисперсию заменить ее приближенным значением

$$\sigma^2\approx s_n^2=\frac{\sum_{k=1}^{n}(x_k-\bar{x})^2}{n-1}$$

(см. главу IV, § 16), то надежность оценки (6.7) уменьшится. Оказывается, что и при неизвестной дисперсии σ^2 можно дать точную оценку приближенного равенства $a\approx\bar{x}$, если исходить не из распределения величины $\bar{\xi}-a$ (которое зависит от σ^2), а из распределения другой случайной величины

$$\zeta=\frac{\bar{\xi}-a}{\frac{1}{\sqrt{n}\sqrt{n-1}}\sqrt{\sum_{k=1}^{n}(\xi_k-\bar{\xi})^2}} \qquad (n\geqslant 2).$$

Если все ξ_k $(k=1,\ 2,\ \ldots,\ n)$ независимы и имеют одно и то же нормальное распределение вероятностей с центром a, то случайная величина ζ имеет распределение, называемое

распределением Стьюдента, с плотностью

$$S(t;\ n) = B_n \left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}},$$

где

$$B_n = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\,\Gamma\left(\frac{n-1}{2}\right)}$$

(Γ — функция Эйлера) *).

Таким образом, вероятность неравенства $|\zeta| < t$ равна

$$\mathbf{P}\{|\zeta| < t\} = \int_{-t}^{t} S(t;\ n)\,dt.$$

Имея таблицы этого интеграла, мы можем по заданной вероятности P найти соответствующее значение $t = t(P,\ n)$, удовлетворяющее уравнению $\int_{-t}^{t} S(t;\ n)\,dt = P$. При этом неравенство $|\zeta| < t$ будет иметь заданную вероятность P.

Переходя снова к опытным значениям x_k величин ξ_k и замечая, что опытным значением величины $\frac{\sum(\xi_k - \overline{\xi})^2}{n-1}$ является как раз выборочная дисперсия s_n^2, мы получаем искомую оценку

$$\left|\frac{\overline{x} - a}{\frac{1}{\sqrt{n}}\,s_n}\right| < t = t(P;\ n)$$

или

$$\overline{x} - t\frac{s_n}{\sqrt{n}} < a < \overline{x} + t\frac{s_n}{\sqrt{n}} \quad \text{с надежностью } P. \qquad (6.8)$$

---

*) Вывод распределения Стьюдента см. Б. В. Гнеденко, Курс теории вероятностей, ГИТТЛ, 1954, § 23. Число величин (n) считается здесь фиксированным.

Ниже приводится таблица *) значений $t = t(P, n)$ для различных значений количества измерений n и обычно задаваемых значений надежности P.

n \ P	0,95	0,99	0,999	n \ P	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,729	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3 374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

В последней строке таблицы (при $n = \infty$) даются значения, совпадающие с соответствующими значениями из таблицы интеграла вероятностей:

$$\Phi(1{,}960) = 0{,}95, \quad \Phi(2{,}576) = 0{,}99, \quad \Phi(3{,}291) = 0{,}999.$$

Объясняется это тем, что при $n \to \infty$ распределение Стьюдента стремится к нормальному распределению, как видно из предельного соотношения

$$\lim_{n \to \infty} \left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}} = e^{-\frac{t^2}{2}}.$$

Расчет средних

Для применения оценки (6.8) надо по результатам измерения вычислить среднее значение $\bar{x} = \frac{\sum x_k}{n}$ и выборочный стан-

*) Эта таблица заимствована из книги: Н. Арлей и К. Бух, Введение в теорию вероятностей и математическую статистику, ИЛ, 1951, стр. 227.

дарт $s_n = \sqrt{\frac{\sum (x_k - \bar{x})^2}{n-1}}$. Непосредственный расчет $\bar{x}$ и s_n по указанным формулам часто оказывается весьма громоздким. Иногда можно существенно упростить расчет с помощью подходящего линейного преобразования результатов измерения:

$$x_k = c + hu_k; \quad u_k = \frac{x_k - c}{h} \qquad (k = 1,\ 2,\ \ldots,\ n). \quad (6.9)$$

За начало отсчета c выбирается некоторое среднее значение между наименьшим и наибольшим значениями x_k; единицу масштаба h выбирают так, чтобы значения u_k выражались целыми числами (что всегда возможно, так как результаты измерения округляются до определенного разряда, связанного с единицей шкалы прибора). Произведя замену (6.9), мы получим следующие расчетные формулы:

$$\bar{x} = c + h\bar{u}, \qquad \text{где } \bar{u} = \frac{\sum u_k}{n}; \qquad (6.10)$$

$$\sum (x_k - \bar{x})^2 = \sum (hu_k - h\bar{u})^2 = h^2 \left(\sum u_k^2 - 2\bar{u} \sum u_k + n\bar{u}^2\right) = h^2 \left(\sum u_k^2 - n\bar{u}^2\right)$$

и, значит,

$$s_n = h \sqrt{\frac{\sum u_k^2 - n\bar{u}^2}{n-1}}. \qquad (6.11)$$

Пример.
Таблица

4,781	4,775	4,764	4,789
4,795	4,772	4,776	4,764
4,769	4,791	4,771	4,774
4,792	4,782	4,789	4,778
4,779	4,767	4,772	4,791

дает результаты произведенных Милликеном первых 20 измерений заряда электрона в 10^{-10} абс. эл.-ст. единиц. Для обработки этих результатов составим расчетную таблицу, выбрав начало отсчета $c = 4{,}780$ и единицу масштаба $h = 0{,}001$ (при этом для удобства расчета мы располагаем результаты в порядке их возрастания). Для расчета нам нужны два столбца — столбец приведенных отклонений (u) и столбец их квадратов (u^2).

Расчетная таблица

x	$u=\frac{x-4{,}780}{0{,}001}$	u^2
4,764	−16	256
4,764	−16	256
4,767	−13	169
4,769	−11	121
4,771	−9	81
4,772	−8	64
4,772	−8	64
4,774	−6	36
4,775	−5	25
4,776	−4	16
4,778	−2	4
4,779	−1	1
4,781	1	1
4,782	2	4
4,789	9	81
4,789	9	81
4,791	11	121
4,791	11	121
4,792	12	144
4,795	15	225
Суммы	−29	1871

Сумма чисел каждого столбца указана в последней строке

$$\sum u_k = -29;$$

$$\sum u_k^2 = 1871.$$

Применяя формулы (6.10) и (6.11), получаем

$$\bar{u} = \frac{-29}{20} = -1{,}45;$$

$$\bar{x} = 4{,}780 - 0{,}00145 = 4{,}77855;$$

$$n(\bar{u})^2 = 20\left(\frac{29}{20}\right)^2 = 42{,}0;$$

$$s_n = 0{,}001\sqrt{\frac{1871-42}{19}} = 0{,}00981.$$

Истинный заряд электрона e, выраженный в 10^{-10} абс. эл.-ст. единиц, мы можем считать приближенно равным

$$e \approx 4{,}7786.$$

Оценим это приближенное равенство, задаваясь надежностью вывода $P=0{,}99$. По таблице на стр. 122 для $P=0{,}99$ и $n=20$ находим $t=2{,}861$.

Следовательно, с надежностью вывода $P=0{,}99$ мы можем утверждать, что истинный заряд заключен между пределами

$$\bar{x} - t\frac{s_n}{\sqrt{n}} = 4{,}77855 - 2{,}861\frac{0{,}00981}{\sqrt{20}} = 4{,}7722$$

и

$$\bar{x} + t\frac{s_n}{\sqrt{n}} = 4{,}77855 + 2{,}861\frac{0{,}00981}{\sqrt{20}} = 4{,}7848,$$

то есть что

$$4{,}7722 < e < 4{,}7848 \qquad (P = 0{,}99).$$

Если увеличить надежность вывода до $P = 0{,}999$, то соответствующее значение t увеличится до 3,883 и, следовательно, увеличится интервал для e:

$$4{,}7700 < e < 4{,}7870 \qquad (P = 0{,}999).$$

Для уменьшения этого интервала надо увеличивать число измерений или повышать точность отдельных измерений.

Оценка точности прибора (точности измерений)

Точность измерений характеризуется величиной σ — стандартом распределения случайных ошибок измерения.

Приближенным значением σ является «выборочный стандарт» $s_n = \sqrt{\frac{\sum (x_k - \bar{x})^2}{n-1}}$. Для оценки приближенного равенства $\sigma \approx s_n$ можно воспользоваться тем, что распределение вероятностей случайной величины

$$\chi = \frac{1}{\sigma}\sqrt{\sum_{k=1}^{n} (\xi_k - \bar{\xi})^2}$$

зависит только от n и не зависит ни от a, ни от σ. Если все ξ_k $(k = 1, 2, \ldots, n)$ независимы и имеют одно и то же нормальное распределение с центром a и дисперсией σ^2, то распределение величины χ имеет плотность

$$R(t;\, n) = A_n t^{n-2} e^{-\frac{t^2}{2}} \qquad (n \geqslant 3;\ t \geqslant 0),$$

где $A_n = \frac{1}{2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)}$ (Γ — функция Эйлера)*). Таким образом, вероятность неравенства $t_1 < \chi < t_2$ равна

$$\mathbf{P}\{t_1 < \chi < t_2\} = \int_{t_1}^{t_2} R(t;\, n)\, dt. \qquad (6.12)$$

---

*) Вывод распределения величины χ см. Б. В. Гнеденко, Курс теории вероятностей, ГИТТЛ, 1954, § 23 и § 64.

Интеграл (6.12) позволяет найти вероятность интересующего нас неравенства $s_n - \varepsilon < \sigma < s_n + \varepsilon$, которое мы будем записывать в виде

$$s_n(1-q) < \sigma < s_n(1+q) \tag{6.13}$$

$\left(\text{обозначая через } q = \frac{\varepsilon}{s_n} \text{ относительную ошибку}\right)$.

Действительно, преобразуем неравенство

$$(1-q)\sqrt{\frac{\sum(\xi_k-\bar{\xi})^2}{n-1}} < \sigma < (1+q)\sqrt{\frac{\sum(\xi_k-\bar{\xi})^2}{n-1}} \tag{6.14}$$

в равносильное и потому равновероятное ему неравенство

$$\frac{\sqrt{n-1}}{1+q} < \frac{1}{\sigma}\sqrt{\sum(\xi_k-\bar{\xi})^2} < \frac{\sqrt{n-1}}{1-q} \quad (q<1).$$

Последнее неравенство имеет вид $t_1 < \chi < t_2$, и поэтому его вероятность равна интегралу (6.12), где

$$t_1 = \frac{\sqrt{n-1}}{1+q}; \quad t_2 = \frac{\sqrt{n-1}}{1-q}.$$

Задаваясь теперь определенной вероятностью P (надежность оценки), мы можем найти соответствующее значение $q = q(P;\, n)$ из уравнения

$$\int\limits_{\frac{\sqrt{n-1}}{1+q}}^{\frac{\sqrt{n-1}}{1-q}} R(t;\, n)\, dt = P,$$

и тогда неравенство (6.14) будет иметь заданную вероятность P. Переходя к опытным значениям x_k случайных величин ξ_k, мы получаем оценку (6.13) с заданной надежностью P.

Замечание. Если $q > 1$, то ввиду положительности стандарта σ неравенство (6.14) примет вид

$$0 < \sigma < (1+q)\sqrt{\frac{\sum(\xi_k-\bar{\xi})^2}{n-1}},$$

что равносильно неравенству

$$t_1 < \chi < +\infty, \quad t_1 = \frac{\sqrt{n-1}}{1+q}.$$

Вероятность P этого неравенства равна

$$P = \int_{\frac{\sqrt{n-1}}{1+q}}^{\infty} R(t;\, n)\, dt.$$

Из этого последнего соотношения и находится $q = q(P;\, n)$ в случае $q > 1$. Следовательно, при $q > 1$ оценка (6.13) принимает вид

$$0 < \sigma < s_n (1 + q).$$

Ниже приводится таблица *) значений $q = q(P;\, n)$ для различных значений количества измерений n и обычно задаваемых значений надежности P.

n \ P	0,95	0,99	0,999	n \ P	0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,221
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Пример.

Для рассмотренных на стр. 124 результатов 20 измерений заряда электрона нами уже было вычислено значение $s_n = 0{,}00981$. Поэтому точность рассматриваемых измерений характеризуется стандартом:

$$\sigma \approx 0{,}00981.$$

*) Эта таблица заимствована из книги В. И. Романовского, Основные задачи теории ошибок, Гостехиздат, 1947.

Оценим это приближенное равенство с надежностью $P=0,99$. По таблице на стр. 127 при $n=20$ и $P=0,99$ находим $q=0,58$. Поэтому можно утверждать с надежностью $P=0,99$, что стандарт ошибок измерения заключен между числами

$$s_n(1-q)=0,00981\,(1-0,58)=0,0041$$

и

$$s_n(1+q)=0,00981\,(1+0,58)=0,0145,$$

то есть что

$$0,0041<\sigma<0,0145 \qquad (P=0,99).$$

Если надежность вывода увеличить до $P=0,999$, то соответствующее значение q увеличится до 0,88 и, следовательно, увеличится интервал для σ:

$$0,0012<\sigma<0,0185 \qquad (P=0,999).$$

Для уменьшения этого интервала надо значительно увеличить число измерений. Например, для нахождения стандарта σ с относительной точностью до 10% надо произвести 350 измерений (при $P=0,99$) или даже 600 измерений (при $P=0,999$).

Упрощенные оценки, «правило трех сигма»

Оценки рассмотренного выше типа требуют изучения специальных распределений (распределение Стьюдента, χ-распределение и т. п.). На практике часто применяют упрощенные оценки, например «правило трех сигма». Это правило заключается в том, что ошибка приближенного равенства $a\approx\overline{x}$ не превосходит утроенной средней квадратической ошибки средней $\overline{x}$. Если величина σ известна, то «правило трех сигма»

$$|a-\overline{x}|<3\sigma(\overline{x})=3\frac{\sigma}{\sqrt{n}}$$

имеет надежность $P=\Phi(3)=0,997$, как видно из оценки (6.7). Но «правило трех сигма» применяют и при неизвестной σ, заменяя ее приближенным значением s_n. «Правило трех сигма» принимает вид

$$|a-\overline{x}|<3\frac{s_n}{\sqrt{n}}, \qquad (6.15)$$

и его надежность оказывается значительно меньшей, чем 0,997, причем эта надежность уменьшается с уменьшением числа измерений n.

Действительно, сравнивая оценку (6.15) с оценкой (6.8), мы замечаем, что при $n=14$ оценка (6.15) имеет надежность, меньшую, чем 0,99, ибо t (0,99; 14) $=3{,}01>3$; а при $n=8$ надежность оценки (6.15) становится равной 0,98.

«Правило трех сигма» применяют также и для оценки некоторых других характеристик распределения, так как нахождение средней квадратической ошибки всегда проще, чем изучение соответствующего распределения. В качестве примера рассмотрим еще применение «правила трех сигма» для оценки стандарта ошибок измерения. Можно подсчитать, что средняя квадратическая ошибка выборочного стандарта приближенно равна

$$\sigma\left(\sqrt{\frac{\sum(\xi_k-\bar{\xi})^2}{n-1}}\right)\approx\frac{s_n}{\sqrt{2(n-1)}}.$$

Поэтому «правило трех сигма» принимает вид

$$|\sigma-s_n|<3\frac{s_n}{\sqrt{2(n-1)}}. \qquad (6.16)$$

Сравнение этой оценки с оценкой (6.13) показывает, что даже при $n=45$ оценка (6.16) имеет надежность, меньшую чем 0,99, ибо q (0,99; 45) $=0{,}321>\frac{3}{\sqrt{2(45-1)}}$. При $n=19$ оценка (6.16) имеет надежность только 0,98, а при $n=7$ ее надежность становится меньше чем 0,95.

Для повышения надежности «правила трех сигма» надо иметь большое число измерений n. При оценке σ этого можно достичь, учитывая измерения одним и тем же прибором различных величин.

Если $n_1, n_2, \ldots, n_m$ — количества измерений первой, второй, ..., m-й величины, а $s_1, s_2, \ldots, s_m$ — соответствующие выборочные стандарты, то «правило трех сигма» принимает вид *)

$$|\sigma-S|<3\frac{S}{\sqrt{2(n-m)}},$$

*) См., например, Н. Арлей и К. Бух, Введение в теорию вероятностей и математическую статистику, ИЛ, 1951. Там же можно найти упрощенные оценки и для функций от измеряемых величин.

где

$$S=\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2+\ldots+(n_m-1)s_m^2}{n-m}}$$

и $n=n_1+n_2+\ldots+n_m$. Надежность этой оценки при $n-m=200$ достигает 0,995.

Упражнения

1. Оценить с надежностью $P=0{,}999$ истинное значение e заряда электрона по результатам 58 измерений Милликена, приведенным в следующей таблице (все данные в 10^{-10} абс. эл.-ст. единиц):

4,781	4,764	4,777	4,809	4,761	4,769
4,795	4,776	4,765	4,790	4,792	4,806
4,769	4,771	4,785	4,779	4,758	4,779
4,792	4,789	4,805	4,788	4,764	4,785
4,779	4,772	4,768	4,772	4,810	4,790
4,775	4,789	4,801	4,791	4,799	4,777
4,772	4,764	4,785	4,788	4,799	4,749
4,791	4,774	4,783	4,783	4,797	4,781
4,782	4,778	4,808	4,740	4,790	
4,767	4,791	4,771	4,775	4,747	

Ответ. Средний результат измерений

$$\overline{x}=4{,}780+0{,}001\cdot 0{,}81=4{,}78081\approx e;$$

$$\frac{s_n}{\sqrt{n}}=0{,}001\cdot\sqrt{\frac{13\,367-38}{57\cdot 58}}=0{,}00201.$$

Оценка $|e-\overline{x}|<t\dfrac{s_n}{\sqrt{n}}$;

$$t=t\,(0{,}999;\ 58)=3{,}470;\ \varepsilon=t\frac{s_n}{\sqrt{n}}=0{,}00697;$$

$$4{,}7738<e<4{,}7878.$$

2. Оценить с надежностью $P=0{,}99$ истинное значение измеряемой величины и точность измерений по следующей таблице результатов 100 измерений:

Результат измерения x	3,18	3,20	3,22	3,24	3,26	3,28	
Частота m	4	18	33	35	9	1	100

Указание. При расчете средней $\bar{x}$ и выборочной дисперсии s_n^2 учитывать частоту каждого результата измерения. С этой целью суммирование производить в столбцах «mu» и «mu^2», где $u = \frac{x - x_0}{h}$. Удобно взять $x_0 = 3{,}22$; $h = 0{,}02$.

Ответ.

$$\bar{x} = 3{,}22 + 0{,}02 \cdot 0{,}3 = 3{,}226 \approx a;$$

$$s_n = 0{,}02 \sqrt{\frac{114 - 9}{99}} = 0{,}0206 \approx \sigma.$$

Оценка истинного значения (a) измеряемой величины

$$t\,(0{,}99;\ 100) = 2{,}627;\quad \varepsilon = t\frac{s_n}{\sqrt{n}} = 0{,}0054;$$

$$|a - 3{,}226| < 0{,}0054 \quad \text{или} \quad 3{,}221 < a < 3{,}232$$

Оценка стандарта (σ) распределения случайных ошибок измерения:

$$0{,}0206\,(1 - q) < \sigma < 0{,}0206\,(1 + q),$$

где

$$q\,(0{,}99;\ 100) = 0{,}198;$$

или

$$0{,}0165 < \sigma < 0{,}0247.$$

3. Оценить с помощью «правила трех сигма» точность измерений по следующим данным десяти серий измерений одним и тем же прибором различных величин:

1-я серия	2-я серия	3-я серия	4-я серия	5-я серия	6-я серия	7-я серия	8-я серия	9-я серия	10-я серия
4,16	6,04	5,71	4,94	4,50	3,84	5,56	5,33	7,89	4,23
4,19	6,05	5,71	4,92	4,48	3,84	5,54	5,32	7,88	4,21
4,15	6,06	5,70	4,93	4,51	3,83	5,56	5,31	7,89	4,22
4,17	6,03	5,71	4,95	4,49	3,84	5,56	5,34	7,87	4,20
4,18	6,05	5,69	4,94	4,52	3,82	5,55	5,32	7,90	4,22
4,17	6,05	5,71	4,93	4,50	3,84	5,57	5,33	7,89	4,21
4,17	6,05	5,71	4,94	4,50	3,85	5,56	5,32	7,91	4,20
4,16	6,06	5,69	4,94	4,50	3,84	5,55	5,35	7,88	4,21
4,17	6,05	5,71	4,93	4,49	3,84	5,56	5,33	7,89	4,22
4,17	6,04	5,73	4,95	4,50	3,86	5,56	5,34	7,90	4,21
4,17	6,05	5,71	4,96	4,49	3,83	5,57	5,33	7,89	4,21
4,16	6,05	5,72	4,94	4,50	3,84	5,56	5,34	7,90	4,19
4,17	6,07	5,72	4,95	4,50	3,83	5,58	5,33	7,89	4,21
4,18	6,06	5,70	4,93	4,51	3,85	5,56	5,34	7,91	4,22
4,17	6,05	5,71	4,94	4,49	3,82	5,58	5,33	7,89	4,21

Указание. При расчете каждой выборочной дисперсии учитывать частоту результатов измерения, например, для обработки первой серии измерений составить расчетную таблицу так:

x	m	$u=\frac{x-4,17}{0,01}$	mu	mu^2
4,15	1	−2	−2	4
4,16	3	−1	−3	3
4,17	8	0	0	0
4,18	2	1	2	2
4,19	1	2	2	4
—	15	—	−1	13

Ответ.

$$14s_1^2=12,93\cdot 10^{-4};\qquad 14s_2^2=12,93\cdot 10^{-4};$$
$$14s_3^2=15,73\cdot 10^{-4};\qquad 14s_4^2=14,93\cdot 10^{-4};$$
$$14s_5^2=13,73\cdot 10^{-4};\qquad 14s_6^2=16,40\cdot 10^{-4};$$
$$14s_7^2=15,73\cdot 10^{-4};\qquad 14s_8^2=14,93\cdot 10^{-4};$$
$$14s_9^2=16,40\cdot 10^{-4};\qquad 14s_{10}^2=13,73\cdot 10^{-4};$$

$$S=\sqrt{\frac{14(s_1^2+s_2^2+\ldots+s_{10}^2)}{140}}=0,01\sqrt{\frac{147,5}{140}}=0,01026\approx\sigma.$$

Оценка

$$|\sigma-S|<3\frac{S}{\sqrt{2\cdot 140}}=0,00184,$$

или

$$0,0084<\sigma<0,0121.$$

ГЛАВА VII

ЛИНЕЙНАЯ КОРРЕЛЯЦИЯ

§ 22. О различных типах зависимостей

Наиболее простым видом связи между величинами является функциональная зависимость, когда каждому значению одной величины соответствует вполне определенное значение другой величины. Такова, например, связь между давлением и объемом газа в сосуде при неизменной температуре. Если в задаче надо учесть изменение давления при одновременном изменении и объема и температуры, то пользуются понятием функции нескольких переменных; при этом снова предполагается, что каждой совокупности значений независимых переменных соответствует вполне определенное значение функции. Функциональные зависимости изучаются в математическом анализе.

Но существуют и такие связи между физическими величинами, которые нельзя отнести к типу функциональных зависимостей. Такова, например, связь между осадками и урожаем или связь между толщиной снегового покрова зимой и объемом стока последующего половодья. Здесь каждому значению одной величины соответствует множество возможных значений другой величины. Рассеяние этих возможных значений объясняется влиянием весьма большого количества дополнительных факторов, от которых мы отвлекаемся, изучая связь между данными величинами. При этом на практике чаще всего ограничиваются изучением изменения средних характеристик одной величины при изменении другой. Поясним это следующим схематическим примером. Пусть 20 опытов над величинами x и y дали результаты, представленные на рис. 19 кружочками.

Изменение величины y при изменении величины x можно характеризовать ломаной, соединяющей средние значения

величины y для каждого значения величины x $\left(\right.$например, при $x=1$ имеем три значения $y=1$; 2; 4; среднее значение $\left.\frac{1+2+4}{3}=2\frac{1}{3}\right)$. Зависимость полученных средних от величины x является уже функциональной — *каждому значению x соответствует вполне определенное среднее значение величины y* *).

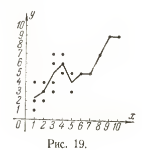


Рис. 19.

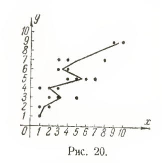


Рис. 20.

В применении к случайным величинам описанный выше тип зависимости приводит к понятию корреляции.

Определение. *Две случайные величины ξ и η находятся в корреляционной зависимости, если каждому значению любой из этих величин соответствует определенное распределение вероятностей другой величины.* Указанные распределения вероятностей называются *условными.*

В настоящей главе мы ограничимся изучением различных средних значений для условных распределений вероятностей; в частности, мы подробно рассмотрим центры условных распределений.

---

*) Полезно обратить внимание на то, что средняя зависимость величины x от величины y приводит уже к другой функциональной зависимости. Так, на рис. 20 для того же примера, что и на рис. 19, ломаная соединяет средние значения величины x для каждого значения величины y $\left(\right.$например, при $y=5$ имеем четыре значения $x=4$; 5; 6; 7; среднее значение $\left.\frac{4+5+6+7}{4}=5{,}5\right)$; при этом две ломаные (на рис. 19 и на рис. 20) заметно отличаются одна от другой.

§ 23. Условные математические ожидания и их свойства

Центр условного распределения величины η при значении $\xi=x$ или условное математическое ожидание величины η при $\xi=x$ определяется как сумма произведений возможных значений величины η на их условные вероятности:

$$\mathbf{M}_x\eta=\sum_y y\mathbf{P}\{\eta=y \mid \xi=x\}, \tag{7.1}$$

где $\mathbf{P}\{\eta=y \mid \xi=x\}$ есть условная вероятность равенства $\eta=y$ при условии, что $\xi=x$, а сумма $\sum\limits_y$ берется по всем значениям y величины η. Для непрерывных распределений эта сумма заменяется соответственно интегралом

$$\mathbf{M}_x\eta=\int\limits_{-\infty}^{\infty} y\varphi_x(y)\,dy, \tag{7.2}$$

где $\varphi_x(y)$ — плотность условного распределения вероятностей величины η при условии, что $\xi=x$.

Условное математическое ожидание $\mathbf{M}_x\eta$ есть функция от x. Эта функция называется *функцией регрессии величины η на величину ξ*; обозначим ее через $f(x)$:

$$f(x)=\mathbf{M}_x\eta.$$

Уравнение $y=f(x)$ называется уравнением регрессии η на ξ, а соответствующая линия — линией регрессии η на ξ.

Нахождение и изучение функции регрессии является одной из основных задач анализа корреляционной зависимости. При решении этой задачи для рассматриваемого далее случая линейной корреляции важную роль играют формулы

$$\mathbf{M}\eta=\mathbf{M}f(\xi), \tag{7.3}$$

$$\mathbf{M}\xi\eta=\mathbf{M}\xi f(\xi). \tag{7.4}$$

Последнюю формулу можно рассматривать как обобщение теоремы умножения математических ожиданий на зависимые случайные величины. Действительно, применяя общее правило

умножения вероятностей (1.13), получаем:

$$M\xi\eta = \sum_{x,y} (xy) P \left\{ \begin{matrix} \xi = x \\ \text{и} \\ \eta = y \end{matrix} \right\} = \sum_x \sum_y xy P\{\xi = x\} P\{\eta = y \mid \xi = x\},$$

где суммы берутся по всем возможным значениям x и y величин ξ и η соответственно. Производя внутреннее суммирование по y, мы можем вынести за знак суммы $\sum\limits_y$ множитель $xP\{\xi = x\}$, не зависящий от y. Это дает

$$M\xi\eta = \sum_x xP\{\xi = x\} \sum_y yP\{\eta = y \mid \xi = x\},$$

то есть

$$M\xi\eta = \sum_x xP\{\xi = x\} M_x\eta = \sum_x xf(x) P\{\xi = x\}, \qquad (7.5)$$

что совпадает с $M\xi f(\xi)$ (см. § 10) *).

Формулы (7.3) и (7.4) являются частными случаями более общей формулы

$$Mu(\xi)\eta = Mu(\xi) f(\xi), \qquad (7.6)$$

где $u(\xi)$ — любая функция, для которой существует $Mu(\xi)\eta$.

Приведем доказательство формулы (7.6) для того случая, когда задана плотность $p(x; y)$ двумерного распределения величины $(\xi; \eta)$. Прежде всего, вероятность попадания в прямоугольник

$$\begin{bmatrix} x < \xi < x + dx \\ y < \eta < y + dy \end{bmatrix}$$

может быть представлена как вероятность совмещения случайных событий $(x < \xi < x + dx)$ и $(y < \eta < y + dy)$. Общее правило

*) Заметим, что если величины ξ и η независимы, то для всех x тождественно выполняется равенство $P\{\eta = y \mid \xi = x\} = P\{\eta = y\}$ и, значит, $M_x\eta = M\eta$. Поэтому для независимых величин формула (7.4) переходит в более простую формулу (3.9). Это же положение будет иметь место и в более общем случае, когда функция $f(x)$ постоянна. Действительно, если $f(x) = b$, то по формуле (7.3) $M\eta = Mb = b$ и по формуле (7.5):

$$M\xi\eta = \sum_x xbP\{\xi = x\} = bM\xi = M\xi M\eta.$$

умножения вероятностей приводит здесь к соотношению между дифференциалами:

$$p(x;\ y)\,dx\,dy = \psi_1(x)\,dx\,\varphi_x(y)\,dy, \qquad (7.7)$$

где $\psi_1(x) = \int\limits_{-\infty}^{\infty} p(x;\ y)\,dy$ есть плотность распределения величины ξ, а $\varphi_x(y)\,dy$ есть дифференциал условной вероятности попадания величины η в интервал $(y,\ y+dy)$ при условии, что $\xi = x$. Из формулы (7.7) видно, что плотность условного распределения вероятностей $\varphi_x(y)$ может быть выражена через плотности $p(x;\ y)$ и $\psi_1(x)$ так:

$$\varphi_x(y) = \frac{p(x;\ y)}{\psi_1(x)}.$$

Подставляя это выражение в формулу (7.2), получаем:

$$f(x) = \mathbf{M}_x\eta = \int\limits_{-\infty}^{\infty} y\,\frac{p(x;\ y)}{\psi_1(x)}\,dy.$$

Теперь простой расчет приводит к формуле (7.6):

$$\mathbf{M}u(\xi)f(\xi) = \int\limits_{-\infty}^{\infty} [u(x)f(x)]\,\psi_1(x)\,dx =$$
$$= \int\limits_{-\infty}^{\infty} u(x)\,\psi_1(x)\left[\int\limits_{-\infty}^{\infty} y\,\frac{p(x;\ y)}{\psi_1(x)}\,dy\right]dx =$$
$$= \iint [u(x)\,y]\,p(x;\ y)\,dx\,dy = \mathbf{M}u(\xi)\,\eta.$$

Формула (7.6) показывает, что *среднее значение любой функции* $u(\xi)\,\eta + v(\xi)$, *линейной относительно* η, *не изменяется при замене величины* η *на функцию регрессии* f *от* ξ.

Отсюда вытекает еще одно важное свойство функции регрессии: *среднее квадратическое отклонение величины* η *от функции* $f(\xi)$ *меньше, чем среднее квадратическое отклонение ее от любой другой функции* $h(\xi)$:

$$\sqrt{\mathbf{M}\,[\eta - f(\xi)]^2} \leqslant \sqrt{\mathbf{M}\,[\eta - h(\xi)]^2}. \qquad (7.8)$$

Для доказательства этого утверждения обозначим $h(\xi) - f(\xi) = u(\xi)$ и воспользуемся свойством линейности математического ожидания:

$$\mathbf{M}\,[\eta - h(\xi)]^2 = \mathbf{M}\,[\eta - f(\xi) - u(\xi)]^2 =$$
$$= \mathbf{M}\,[\eta - f(\xi)]^2 + \mathbf{M}\,[u(\xi)]^2 - 2\mathbf{M}\,[\eta - f(\xi)]\,u(\xi).$$

А так как в силу формулы (7.6) последнее слагаемое равно нулю

$$\mathbf{M}\,[\eta - f(\xi)]\,u(\xi) = \mathbf{M}\eta u(\xi) - \mathbf{M}f(\xi)\,u(\xi) = 0,$$

то

$$\mathbf{M}\,[\eta - h(\xi)]^2 = \mathbf{M}\,[\eta - f(\xi)]^2 + \mathbf{M}\,[h(\xi) - f(\xi)]^2, \qquad (7.9)$$

откуда и вытекает неравенство (7.8).

Подчеркнем, что свойство (7.6) относится только к линейным относительно η функциям; средние значения нелинейных функций могут измениться при замене η на $f(\xi)$; например, для дисперсий имеет место неравенство

$$\sigma^2(\eta) = \mathbf{M}\,(\eta - \mathbf{M}\eta)^2 \geqslant \sigma^2\,[f(\xi)]. \qquad (7.10)$$

Это неравенство легко получить из соотношения (7.9), полагая в нем

$$h(\xi) = b = \mathbf{M}\eta = \mathbf{M}f(\xi).$$

При этом

$$\mathbf{M}\,(\eta - b)^2 = \mathbf{M}\,[\eta - f(\xi)]^2 + \mathbf{M}\,[f(\xi) - b]^2 \geqslant \sigma^2\,[f(\xi)].$$

Выше мы определили функцию регрессии η на ξ. Совершенно аналогично определяется *функция регрессии* ξ *на* η:

$$\mathbf{M}_y\xi = \sum_x x\mathbf{P}\,\{\xi = x \mid \eta = y\} = g(y).$$

Следует иметь в виду, что если связь между ξ и η не является строго функциональной, то функции $f(x)$ и $g(y)$ не являются взаимно обратными и, значит, линии регрессии η на ξ и ξ на η не совпадают.

§ 24. Линейная корреляция

Определение. *Корреляционная зависимость между случайными величинами* ξ *и* η *называется линейной корреляцией, если обе функции регрессии* $f(x)$ *и* $g(y)$ *являются линейными.* В этом случае обе линии регрессии являются прямыми; они называются *прямыми регрессии*.

Выведем уравнение прямой регрессии η на ξ, то есть найдем коэффициенты линейной функции

$$f(x) = Ax + B.$$

Обозначим $\mathbf{M}\xi = a$; $\mathbf{M}\eta = b$; $\mathbf{M}(\xi - a)^2 = \sigma_1^2$; $\mathbf{M}(\eta - b)^2 = \sigma_2^2$. Прежде всего, с помощью формулы (7.3) находим:

$$\mathbf{M}\eta = \mathbf{M}f(\xi) = \mathbf{M}\,(A\xi + B),$$

то есть

$$b = Aa + B,$$

откуда

$$B = b - Aa.$$

Далее, с помощью формулы (7.4) находим:

$$\mathbf{M}\xi\eta = \mathbf{M}\xi f(\xi) = \mathbf{M}(A\xi^2 + B\xi) = A\mathbf{M}\xi^2 + (b - Aa)\,a,$$

откуда

$$A = \frac{\mathbf{M}\xi\eta - ab}{\mathbf{M}\xi^2 - a^2} = \frac{\mathbf{M}\xi\eta - ab}{\sigma_1^2}.$$

Полученный коэффициент называют *коэффициентом регрессии η на ξ* и обозначают через $\rho(\eta/\xi)$:

$$\rho(\eta/\xi) = \frac{\mathbf{M}\xi\eta - ab}{\sigma_1^2}.$$

Таким образом, уравнение прямой регрессии η на ξ имеет вид

$$y = \rho(\eta/\xi)(x - a) + b. \tag{7.11}$$

Аналогично можно получить уравнение прямой регрессии ξ на η:

$$x = \rho(\xi/\eta)(y - b) + a, \tag{7.12}$$

где

$$\rho(\xi/\eta) = \frac{\mathbf{M}\xi\eta - ab}{\sigma_2^2}$$

есть коэффициент регрессии ξ на η.

Уравнения прямых регрессии можно записать в более симметричном виде, если ввести безразмерный, симметричный относительно ξ и η, коэффициент

$$r = \frac{\mathbf{M}\xi\eta - ab}{\sigma_1\sigma_2}, \tag{7.13}$$

который называется *коэффициентом корреляции между величинами ξ и η*. При этом

$$\rho(\eta/\xi) = r\frac{\sigma_2}{\sigma_1}; \qquad \rho(\xi/\eta) = r\frac{\sigma_1}{\sigma_2}$$

и уравнения прямых регрессии принимают вид

$$\frac{y - b}{\sigma_2} = r\,\frac{x - a}{\sigma_1}, \tag{7.11*}$$

$$\frac{x - a}{\sigma_1} = r\,\frac{y - b}{\sigma_2}. \tag{7.12*}$$

Из уравнений прямых регрессии видно, что обе эти прямые проходят через точку $(a; b)$ — центр совместного распределения величин ξ, η; угловые коэффициенты прямых регрессии равны соответственно (обозначения углов см. на рис. 21)

$$\operatorname{tg}\alpha = r\frac{\sigma_2}{\sigma_1}; \qquad \operatorname{tg}\beta = \frac{1}{r}\frac{\sigma_2}{\sigma_1}.$$

В следующем параграфе мы докажем, что $|r| \leqslant 1$ и, следовательно, $|\operatorname{tg}\alpha| \leqslant |\operatorname{tg}\beta|$. Это означает, что прямая регрессии η на ξ имеет меньший наклон к оси абсцисс, чем прямая регрессии ξ на η. Чем ближе $|r|$ к 1, тем меньше угол между прямыми регрессии. Эти прямые сливаются тогда и только тогда, когда $|r| = 1$.

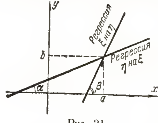


Рис. 21.

При $r = 0$ прямые регрессии имеют уравнения $y = b$; $x = a$. В этом случае $\mathbf{M}_x\eta = b = \mathbf{M}\eta$; $\mathbf{M}_y\xi = a = \mathbf{M}\xi$.

Коэффициенты регрессии имеют тот же знак, что и коэффициент корреляции r, и связаны соотношением

$$\rho(\eta/\xi)\,\rho(\xi/\eta) = r^2. \qquad (7.14)$$

Из того, что знаки $\rho(\eta/\xi)$ и $\rho(\xi/\eta)$ одинаковы, следует, между прочим, что если величина η в среднем возрастает при увеличении величины ξ, то и величина ξ в среднем возрастает при увеличении величины η; но связь между скоростями возрастания этих величин существенно зависит от коэффициента корреляции.

Пример. Нормальная корреляция.
Корреляция между ξ и η называется нормальной, если плотность двумерного распределения вероятностей величины $(\xi; \eta)$ задается формулой

$$p(x; y) = \frac{\sqrt{AC - B^2}}{2\pi} e^{-\frac{1}{2}[A(x-a)^2 + 2B(x-a)(y-b) + C(y-b)^2]},$$

где A, B, C — некоторые постоянные, $A > 0$, $C > 0$, $AC - B^2 > 0$.

При этом плотность частного распределения величины ξ согласно формуле (2.36) будет равна

$$\psi_1(x)=\int_{-\infty}^{\infty} p(x;\ y)\,dy=$$

$$=\frac{\sqrt{AC-B^2}}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}C\left[(y-b)+\frac{B}{C}(x-a)\right]^2-\frac{1}{2}\left(A-\frac{B^2}{C}\right)(x-a)^2}\,dy=$$

$$=\frac{\sqrt{AC-B^2}}{\sqrt{2\pi}\sqrt{C}}\,e^{-\frac{1}{2}\frac{AC-B^2}{C}(x-a)^2},$$

так как

$$\frac{\sqrt{C}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}C(y-\lambda)^2}\,dy=1 \quad \text{при любом } \lambda.$$

Отсюда видно, что частное распределение величины ξ есть нормальное распределение с центром a и дисперсией

$$\sigma_1^2=\frac{C}{AC-B^2}.$$

Аналогично, частное распределение величины η есть нормальное распределение с центром b и дисперсией

$$\sigma_2^2=\frac{A}{AC-B^2}.$$

Условным распределением величины η при фиксированном значении $\xi=x$ также является нормальное распределение с плотностью

$$\varphi_x(y)=\frac{p(x;\ y)}{\psi_1(x)}=\frac{\sqrt{C}}{\sqrt{2\pi}}\,e^{-\frac{1}{2}C\left[(y-b)+\frac{B}{C}(x-a)\right]^2}$$

и, значит, с центром

$$\mathbf{M}_x\eta=b-\frac{B}{C}(x-a).$$

Аналогично находим центр условного распределения величины ξ при фиксированном значении $\eta=y$:

$$\mathbf{M}_y\xi=a-\frac{B}{A}(y-b).$$

Отсюда видно, что нормальная корреляция является линейной корреляцией. Прямые регрессии здесь имеют уравнения

$$y=-\frac{B}{C}(x-a)+b,$$

$$x=-\frac{B}{A}(y-b)+a,$$

и, значит, коэффициенты регрессии равны соответственно

$$\rho(\eta/\xi) = -\frac{B}{C}; \qquad \rho(\xi/\eta) = -\frac{B}{A}.$$

Из последних формул и формулы (7.14) легко найти коэффициент корреляции:

$$r = -\frac{B}{\sqrt{AC}}. \tag{7.15}$$

Подробнее о нормальной корреляции и ее значении см. С. Н. Бернштейн, Теория вероятностей, Гостехиздат (часть 5, главы 3 и 4).

§ 25. Коэффициент корреляции

Рассмотрим подробнее введенный в предыдущем параграфе коэффициент корреляции между случайными величинами ξ и η:

$$r = r(\xi;\ \eta) = \frac{M\xi\eta - ab}{\sigma_1\sigma_2} = \frac{M\xi\eta - M\xi M\eta}{\sigma(\xi)\,\sigma(\eta)}. \tag{7.13*}$$

Этот коэффициент характеризует относительную величину отклонения математического ожидания произведения ($M\xi\eta$) от произведения математических ожиданий ($M\xi M\eta$) величин ξ и η. Так как это отклонение имеет место только для зависимых величин, то можно сказать, что коэффициент корреляции характеризует тесноту зависимости между ξ и η.

С помощью коэффициента корреляции можно обобщить теорему сложения дисперсий для зависимых величин:

$$\sigma^2(\xi + \eta) = \sigma^2(\xi) + \sigma^2(\eta) + 2r(\xi;\ \eta)\,\sigma(\xi)\,\sigma(\eta). \tag{7.16}$$

Эта формула получается из формулы § 12 (см. стр. 72):

$$\sigma^2(\xi + \eta) = M(\xi - a)^2 + 2M(\xi - a)(\eta - b) + M(\eta - b)^2$$

с помощью преобразования

$$M(\xi - a)(\eta - b) =$$
$$= M\xi\eta - aM\eta - bM\xi + ab = M\xi\eta - ab = r\sigma_1\sigma_2.$$

Безразмерность коэффициента корреляции позволяет представить его в виде математического ожидания произведения нормированных отклонений

$$\xi_0 = \frac{\xi - a}{\sigma_1} \quad \text{и} \quad \eta_0 = \frac{\eta - b}{\sigma_2}.$$

Действительно,

$$M\xi_0\eta_0 = M\left(\frac{\xi - a}{\sigma_1}\ \frac{\eta - b}{\sigma_2}\right) = \frac{M(\xi - a)(\eta - b)}{\sigma_1\sigma_2} = r. \quad (7.17)$$

Свойства коэффициента корреляции

Теорема 1. *Линейные преобразования, сводящиеся к изменению масштаба или начала отсчета случайных величин ξ и η, не изменяют коэффициента корреляции между ними:*

$$r(c_1\xi + c_2;\ c_3\eta + c_4) = r(\xi;\ \eta)$$

при любых постоянных $c_1 > 0,\ c_2,\ c_3 > 0,\ c_4$.

Это вытекает из того, что при указанных линейных преобразованиях не изменяются нормированные отклонения ξ_0 и η_0.

Действительно, при замене ξ на $\xi' = c_1\xi + c_2$ $(c_1 > 0)$ имеем:

$$M\xi' = c_1 M\xi + c_2 = c_1 a + c_2;\ \sigma(\xi') = c_1\sigma(\xi) = c_1\sigma_1,$$

и поэтому

$$\xi'_0 = \frac{\xi' - M\xi'}{\sigma(\xi')} = \frac{(c_1\xi + c_2) - (c_1 a + c_2)}{c_1\sigma_1} = \frac{\xi - a}{\sigma_1} = \xi_0.$$

Теорема 2. *Коэффициент корреляции $r(\xi;\ \eta)$ заключен между -1 и $+1$, достигая этих крайних значений только в случае линейной функциональной зависимости между ξ и η.*

Доказательство. Из формулы (7.17) и формул

$$M\xi_0^2 = \frac{M(\xi - a)^2}{\sigma_1^2} = 1; \qquad M\eta_0^2 = 1$$

вытекает равенство

$$M(\xi_0 \pm \eta_0)^2 = M\xi_0^2 \pm 2M\xi_0\eta_0 + M\eta_0^2 = 1 \pm 2r(\xi;\ \eta) + 1.$$

Отсюда следует, что

$$1 \pm r(\xi;\ \eta) = \frac{1}{2} M(\xi_0 \pm \eta_0)^2 \geqslant 0, \quad (7.18)$$

то есть

$$-1 \leqslant r(\xi;\ \eta) \leqslant +1.$$

Знак равенства в соотношении (7.18) достигается тогда и только тогда, когда $M(\xi_0 \pm \eta_0)^2 = 0$, то есть когда $\xi_0 \pm \eta_0 = 0$.

Последнее равенство означает наличие линейной функциональной зависимости

$$\frac{\xi - a}{\sigma_1} \pm \frac{\eta - b}{\sigma_2} = 0,$$

или

$$\eta = b \mp \frac{\sigma_2}{\sigma_1}(\xi - a).$$

Теорема 3. *Коэффициент корреляции между независимыми случайными величинами равен нулю.*

Это непосредственно следует из формул (3.9) и (7.13*).

Заметим, что обратное утверждение неверно, то есть из равенства нулю коэффициента корреляции $r(\xi;\ \eta)$ не следует независимость величин ξ и η. Если коэффициент корреляции

$$r(\xi;\ \eta) = 0,$$

то величины ξ и η называются *некоррелированными*.

В одном важном случае некоррелированность случайных величин ξ, η влечет за собой их независимость. Это имеет место при нормальной корреляции.

Действительно, из формулы (7.15) видно, что при нормальной корреляции коэффициент $r(\xi;\ \eta)$ обращается в нуль тогда и только тогда, когда $B = 0$. Но при этом условии плотность распределения $p(x;\ y)$ можно представить в виде

$$p(x; y) = \frac{\sqrt{AC}}{2\pi} e^{-\frac{1}{2}[A(x-a)^2 + C(y-b)^2]} =$$

$$= \frac{\sqrt{A}}{\sqrt{2\pi}} e^{-\frac{1}{2}A(x-a)^2} \frac{\sqrt{C}}{\sqrt{2\pi}} e^{-\frac{1}{2}C(y-b)^2},$$

откуда сразу видно, что величины ξ и η независимы.

§ 26. Наилучшее линейное приближение к функции регрессии

В рассмотренном нами случае линейной корреляции параметры функции регрессии находятся сравнительно легко. В случае более сложной корреляционной зависимости нахождение функции регрессии представляет значительные трудности. Поэтому возникает задача о наилучшем линейном приближении к функции регрессии. Какой смысл следует вкладывать в слова «наилучшее приближение»? В § 23 мы установили,

что функция $f(\xi)$ дает наилучшее приближение к величине η в смысле среднего квадратического приближения, а именно, что при любой функции $h(\xi)$ имеет место неравенство

$$\mathbf{M}[\eta - f(\xi)]^2 \leqslant \mathbf{M}[\eta - h(\xi)]^2.$$

Естественно теперь под линейной функцией наилучшего приближения к функции регрессии $f(x)$ понимать такую функцию $Ax + B$, для которой средний квадрат отклонения

$$\mathbf{M}[\eta - (A\xi + B)]^2$$

принимает наименьшее значение. Задачей настоящего параграфа является нахождение параметров A и B такой линейной функции. Оказывается, что эти параметры находятся по тем же формулам, что и параметры линейной функции регрессии в случае линейной корреляции (см. § 24). А именно, имеет место следующая

Теорема. *Среднее квадратическое отклонение случайной величины η от линейной функции $A\xi + B$ достигает наименьшего значения тогда и только тогда, когда*

$$A = \rho(\eta/\xi) = r\frac{\sigma_2}{\sigma_1}; \qquad B = b - Aa,$$

то есть когда эта функция имеет вид

$$r\frac{\sigma_2}{\sigma_1}(\xi - a) + b.$$

Другими словами, при любой корреляционной зависимости из всех прямых линий прямая регрессии (7.11) дает наилучшее в среднем приближение к действительной регрессии η на ξ.

Доказательство. Обозначим $B - (b - Aa) = C$ и преобразуем средний квадрат отклонения

$$\mathbf{M}[\eta - (A\xi + B)]^2 = \mathbf{M}[(\eta - b) - A(\xi - a) - C]^2,$$

пользуясь линейностью математического ожидания и учитывая, что $\mathbf{M}(\xi - a) = 0$ и $\mathbf{M}(\eta - b) = 0$:

$$\mathbf{M}[\eta - (A\xi + B)]^2 = \mathbf{M}(\eta - b)^2 + A^2\mathbf{M}(\xi - a)^2 - \\ - 2A\mathbf{M}(\xi - a)(\eta - b) + C^2 = \sigma_2^2 + A^2\sigma_1^2 - 2Ar\sigma_1\sigma_2 + C^2. \quad (7.19)$$

В полученной сумме слагаемое σ_2^2 постоянно, слагаемое C^2 принимает наименьшее значение (0) при $B = b - Aa$, а слагаемое

$$A^2\sigma_1^2 - 2Ar\sigma_1\sigma_2 = \sigma_1^2\left(A - r\frac{\sigma_2}{\sigma_1}\right)^2 - r^2\sigma_2^2$$

принимает наименьшее значение $(-r^2\sigma_2^2)$ при $A = r\frac{\sigma_2}{\sigma_1}$. Тем самым теорема доказана.

Найдем еще средний квадрат отклонения η от линейной функции $r\frac{\sigma_2}{\sigma_1}(\xi - a) + b$, дающей наилучшее линейное приближение в среднем.

Из формулы (7.19) при $C = 0$, $A = r\frac{\sigma_2}{\sigma_1}$ получаем:

$$\mathbf{M}\left\{\eta - \left[r\frac{\sigma_2}{\sigma_1}(\xi - a) + b\right]\right\}^2 = \sigma_2^2 - r^2\sigma_2^2 = \sigma_2^2(1 - r^2). \quad (7.20)$$

Формула (7.20) позволяет выяснить вопрос о том, как именно коэффициент корреляции характеризует тесноту связи.

Действительно, из формулы (7.20) следует, что

$$\sqrt{1 - r^2} = \frac{\sqrt{\mathbf{M}\left\{\eta - \left[r\frac{\sigma_2}{\sigma_1}(\xi - a) + b\right]\right\}^2}}{\sigma_2} = \frac{\sigma\left\{\eta - \left[r\frac{\sigma_2}{\sigma_1}(\xi - a) + b\right]\right\}}{\sigma(\eta)}. \quad (7.21)$$

Отсюда видно, что коэффициент корреляции $r(\xi; \eta)$ характеризует относительную величину среднего квадратического отклонения случайной величины η от линейной функции наилучшего приближения $r\frac{\sigma_2}{\sigma_1}(\xi - a) + b$, т. е. *коэффициент корреляции характеризует тесноту линейной связи между* ξ и η. Чем ближе значение r^2 к 1, тем меньше рассеяны в среднем значения величины η относительно прямой регрессии η на ξ.

Все сказанное выше относится, конечно, и к регрессии ξ на η.

§ 27. Анализ линейной корреляции по данным случайной выборки. Оценка значимости коэффициента корреляции

Для анализа линейной корреляции между двумя величинами ξ и η производят ряд независимых испытаний (опытов, наблюдений), исходом каждого из которых является пара $(x_i;\ y_i)$. Рассматривая полученные пары соответственных значений

$$(x_1;\ y_1),\ (x_2;\ y_2),\ \ldots,\ (x_n;\ y_n)$$

как случайную выборку из совокупности всех возможных значений величины $(\xi;\ \eta)$, мы можем найти приближенные значения всех параметров линейной корреляции между ξ и η по методу моментов (см. главу IV). Прежде всего, мы имеем следующие приближенные формулы:

$$\left.\begin{aligned} a = \mathbf{M}\xi \approx \bar{x} = \frac{\sum x}{n};\ b = \mathbf{M}\eta \approx \bar{y} = \frac{\sum y}{n};\\ \sigma^2(\xi) \approx s_1^2 = \frac{\sum (x-\bar{x})^2}{n-1};\ \sigma^2(\eta) \approx s_2^2 = \frac{\sum (y-\bar{y})^2}{n-1}; \end{aligned}\right\} \quad (7.22)$$

$$\mathbf{M}(\xi - a)(\eta - b) \approx \frac{\sum (x-\bar{x})(y-\bar{y})}{n-1}. \quad (7.23)$$

Отсюда получаем приближенную формулу для коэффициента корреляции

$$r(\xi;\ \eta) \approx r_n = \frac{\sum (x-\bar{x})(y-\bar{y})}{(n-1)s_1 s_2} = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2}\ \sqrt{\sum (y-\bar{y})^2}}. \quad (7.24)$$

Величина r_n называется *выборочным коэффициентом корреляции*.

Далее, заменяя в формулах (7.11) и (7.12) все математические ожидания соответствующими средними значениями, мы получаем *выборочную прямую регрессии* η *на* ξ

$$y - \bar{y} = r_n \frac{s_2}{s_1}(x - \bar{x}) \quad (7.25)$$

и *выборочную прямую регрессии* ξ *на* η

$$x - \bar{x} = r_n \frac{s_1}{s_2}(y - y). \quad (7.26)$$

Коэффициенты $r_n \frac{s_2}{s_1}$ и $r_n \frac{s_1}{s_2}$ называются *выборочными коэффициентами регрессии.*

Важно отметить, что выборочные прямые регрессии (7.25) и (7.26) обладают минимальным свойством, аналогичным рассмотренному в § 26. А именно, сумма квадратов отклонений наблюденных значений y_i от выборочной прямой регрессии (7.25) меньше, чем сумма квадратов отклонений их от любой другой прямой:

$$\sum_{i=1}^{n} \left\{ y_i - \left[\bar{y} + r_n \frac{s_2}{s_1} (x_i - \bar{x}) \right] \right\}^2 \leqslant \sum_{i=1}^{n} \left\{ y_i - (Ax_i + B) \right\}^2.$$

Доказательство проводится тем же методом, что и в § 26. То же самое можно сказать и о прямой регрессии (7.26).

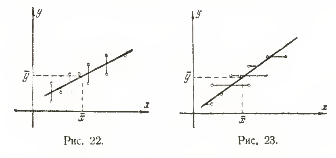


Рис. 22. Рис. 23.

Рисунки 22 и 23 показывают, о каких отклонениях идет здесь речь.

Как видно из приведенных выше формул (7.22) — (7.26), расчет выборочных прямых регрессии связан с большим количеством приближенных вычислений с многозначными числами $(x_i - \bar{x})$ и $(y_i - \bar{y})$. Эти вычисления, как и в § 21, можно существенно упростить с помощью предварительного линейного преобразования величин x и y, то есть с помощью выбора удобного начала отсчета и подходящего масштаба. Заменяя x и y величинами

$$u = \frac{x - x_0}{h_1}; \qquad v = \frac{y - y_0}{h_2} \qquad (h_1 > 0, \ h_2 > 0),$$

мы получим следующие формулы (см. § 21, стр. 123):

$$x = x_0 + h_1 u; \qquad y = y_0 + h_2 v;$$

$$\bar{x} = x_0 + h_1 \bar{u} \left(\bar{u} = \frac{\sum u_i}{n} \right); \qquad \bar{y} = y_0 + h_2 \bar{v} \left(\bar{v} = \frac{\sum v_i}{n} \right);$$

$$s_1 = h_1 \sqrt{\frac{\sum u_i^2 - n(\bar{u})^2}{n-1}}; \qquad s_2 = h_2 \sqrt{\frac{\sum v_i^2 - n(\bar{v})^2}{n-1}}$$

и, наконец,

$$r_n = \frac{\sum (h_1 u_i - h_1 \bar{u})(h_2 v_i - h_2 \bar{v})}{h_1 \sqrt{\sum u_i^2 - n(\bar{u})^2}\, h_2 \sqrt{\sum v_i^2 - n(\bar{v})^2}} = \frac{\sum u_i v_i - n\bar{u}\bar{v}}{\sqrt{\sum u_i^2 - n(\bar{u})^2}\sqrt{\sum v_i^2 - n(\bar{v})^2}}.$$

(Здесь мы заменили

$$\sum (u_i - \bar{u})(v_i - \bar{v}) = \sum u_i v_i - \bar{u} \sum v_i - \bar{v} \sum u_i + n\bar{u}\bar{v} = \sum u_i v_i - n\bar{u}\bar{v}.)$$

Пример расчета. Произведем расчет линейной корреляции по данным следующей таблицы (частота m_i показывает, сколько раз встретилась пара соответственных значений x_i, y_i):

x	y	Частота m
23,0	0,48	2
24,0	0,50	4
24,5	0,49	3
24,5	0,50	2
25,0	0,51	1
25,5	0,52	1
26,0	0,49	2
26,0	0,51	1
26,0	0,53	2
26,5	0,50	1
26,5	0,52	1
27,0	0,54	2
27,0	0,52	1
28,0	0,53	3
		$n = 26$

Выберем для величины x начало отсчета $x_0 = 26{,}0$, масштабный коэффициент $h_1 = 0{,}5$. Для величины y выберем

соответственно $y_0 = 0{,}50$; $h_2 = 0{,}01$. Таким образом, для нашего примера

$$u = \frac{x - 26{,}0}{0{,}5}; \qquad v = \frac{y - 0{,}50}{0{,}01}.$$

Составим расчетную таблицу для вычисления нужных нам сумм $\sum u$, $\sum u^2$, $\sum v$, $\sum v^2$, $\sum uv$, причем заметим, что каждое слагаемое следует учитывать столько раз, сколько раз оно встречается в таблице. Это значит, что соответствующие слагаемые следует еще умножить на частоту m.

Ниже приведена расчетная таблица (во второй строке условно указан порядок действий):

x	y	m	u	um	u^2m	v	vm	v^2m	uvm
(1)	(2)	(3)	(4)	(5) = (3)·(4)	(6) = (4)·(5)	(7)	(8) = (3)·(7)	(9) = (7)·(8)	(10) = (5)·(7)
23,0	0,48	2	—6	—12	72	—2	—4	8	24
24,0	0,50	4	—4	—16	64	0	0	0	0
24,5	0,49	3	—3	—9	27	—1	—3	3	9
24,5	0,50	2	—3	—6	18	0	0	0	0
25,0	0,51	1	—2	—2	4	1	1	1	—2
25,5	0,52	1	—1	—1	1	2	2	4	—2
26,0	0,49	2	0	0	0	—1	—2	2	0
26,0	0,51	1	0	0	0	1	1	1	0
26,0	0,53	2	0	0	0	3	6	18	0
26,5	0,50	1	1	1	1	0	0	0	0
26,5	0,52	1	1	1	1	2	2	4	2
27,0	0,54	2	2	4	8	4	8	32	16
27,0	0,52	1	2	2	4	2	2	4	4
28,0	0,53	3	4	12	48	3	9	27	36
Суммы		26	—	— 26	248	—	22	104	87

В последней строке расчетной таблицы подсчитаны суммы, нужные для дальнейших расчетов *):

$$\sum u = \sum{}^* um = -26; \ \sum u^2 = \sum{}^* u^2m = 248;$$
$$\sum v = \sum{}^* vm = 22; \ \sum v^2 = \sum{}^* v^2m = 104;$$
$$\sum uv = \sum{}^* uvm = 87.$$

*) Здесь через $\sum^*$ обозначены суммы, взятые с учетом частоты m.

Отсюда получаем:

$$\bar{u}=\frac{-26}{26}=-1; \quad \bar{x}=26,0+0,5(-1)=25,5;$$

$$s_1=0,5\sqrt{\frac{248-26(-1)^2}{25}}=1,49;$$

$$\bar{v}=\frac{22}{26}=0,846; \ \bar{y}=0,50+0,01\cdot 0,846=0,50846;$$

$$s_2=0,01\sqrt{\frac{104-26(0,846)^2}{25}}=0,0185;$$

и, наконец, выборочный коэффициент корреляции

$$r_n=\frac{87-26(-1)\,0,846}{\sqrt{248-26}\sqrt{104-18,6}}=\frac{109,0}{14,90\cdot 9,24}=0,793.$$

С помощью полученных данных можно уже написать уравнения прямых регрессии

$$y-0,508=0,793\frac{0,0185}{1,49}(x-25,5)=0,0098(x-25,5),$$

$$x-25,5=0,793\frac{1,49}{0,0185}(y-0,508)=64(y-0,508).$$

На рис. 24 даны прямые регрессии и выборочные данные для рассматриваемого примера.

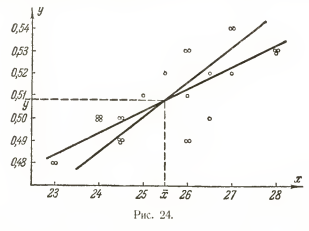


Рис. 24.

Замечание о доверительных оценках коэффициента корреляции.
Рассмотрение вопроса о доверительных оценках коэффициента корреляции выходит за рамки нашего пособия. Заметим только, что

для этих оценок не рекомендуется применять «правило трех сигма», так как распределение вероятностей выборочного коэффициента корреляции даже для больших n значительно отличается от нормального. Мы ограничимся указанием о решении более простого вопроса: не может ли оказаться, что выборочный коэффициент корреляции случайно отличен от нуля, а в действительности случайные величины ξ и η не коррелированы?

Решение этого вопроса дается с помощью распределения вероятностей для выборочного коэффициента корреляции при условии, что истинный коэффициент корреляции $r(\xi;\ \eta)$ равен нулю.

Ниже приводится таблица границ случайных отклонений от нуля произведения $|r_n|$ на $\sqrt{n-1}$ при условии $r(\xi;\ \eta)=0$ в зависимости от заданной вероятности P и числа данных n *). Если для выборочного коэффициента корреляции r_n величина $\sqrt{n-1}\,|r_n|$ окажется больше приведенного в таблице граничного значения, то с надежностью P мы можем утверждать, что истинный коэффициент корреляции $r(\xi;\ \eta)$ отличен от нуля.

Таблица границ случайных отклонений $\sqrt{n-1}\,|r_n|$

n \ P	0,99	0,999	n \ P	0,99	0,999
10	2,29	2,62	25	2,47	3,03
11	2,32	2,68	30	2,49	3,07
12	2,35	2,73	35	2,50	3,10
13	2,37	2,77	40	2,51	3,13
14	2,39	2,81	45	2,52	3,15
15	2,40	2,85	50	2,53	3,16
16	2,41	2,87	60	2,536	3,184
17	2.42	2,90	70	2,541	3,198
18	2,43	2,92	80	2,546	3,209
19	2,44	2,94	90	2,550	3,219
20	2,45	2,96	100	2,553	3,226
			∞	2,576	3,291

Пример. Для примера, рассмотренного на стр. 149—151,

$$n=26;\quad r_n=0{,}793;\quad \sqrt{n-1}\,|r_n|=3{,}96.$$

Так как это число (3,96) значительно больше границы случайного отклонения, которая составляет всего 3,03 при надежности $P=0{,}999$, то мы можем быть уверены в корреляционной связи рассматриваемых величин.

*) И при дополнительном условии, что исследуемая корреляция мало отличается от нормальной. Приведенная здесь таблица заимствована из книги: Н. Арлей и К. Бух, Введение в теорию вероятностей и математическую статистику, ИЛ, Москва, 1951.

Упражнения

1. Найти коэффициент корреляции между величинами λ_1 и λ_2, рассмотренными в упражнении 1 к главе III (стр. 77).

Ответ.

$$r(\lambda_1;\ \lambda_2)=\frac{M\lambda_1\lambda_2-M\lambda_1 M\lambda_2}{\sigma_1\sigma_2}=\frac{\frac{25}{100}\cdot\frac{24}{99}-\left(\frac{25}{100}\right)^2}{\frac{25}{100}\cdot\frac{75}{100}}=-\frac{1}{99}.$$

2. Произвести расчет линейной корреляции по данным следующих выборочных наблюдений:

x	y	Частоты
2,3	7,1	5
2,3	7,3	4
2,6	7,3	12
2,6	7,5	8
2,6	7,7	1
2,9	7,5	5
2,9	7,7	5
3,2	7,5	4
3,2	7,7	7
3,5	7,7	2
3,5	7,9	1
3,8	7,9	1
		55

Ответ.

$$\bar{x}=2{,}9+0{,}3\left(-\frac{19}{55}\right)=2{,}796;\quad \bar{y}=7{,}5+0{,}2\left(-\frac{7}{55}\right)=7{,}475;$$

$$s_1=0{,}3\sqrt{\frac{89-\frac{19^2}{55}}{54}}=0{,}3\cdot\sqrt{\frac{82{,}42}{54}}=0{,}371;$$

$$s_2=0{,}2\sqrt{\frac{59-\frac{7^2}{55}}{54}}=0{,}2\cdot\sqrt{\frac{58{,}11}{54}}=0{,}208;$$

$$r_n=\frac{60-\frac{19\cdot 7}{55}}{\sqrt{82{,}42\cdot 58{,}11}}=\frac{57{,}58}{69{,}21}=0{,}833.$$

Выборочные прямые регрессии

$$y-7{,}47=0{,}833\,\frac{0{,}208}{0{,}371}(x-2{,}80)=0{,}467\,(x-2{,}80),$$

$$x-2{,}80=0{,}833\,\frac{0{,}371}{0{,}208}(y-7{,}47)=1{,}49\,(y-7{,}47).$$

ПРИЛОЖЕНИЕ

ЗНАЧЕНИЯ ИНТЕГРАЛА ВЕРОЯТНОСТЕЙ

$$\Phi(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-\frac{x^2}{2}} dx; \quad \Phi(-t) = -\Phi(t)$$

Таблица I

t	$\Phi(t)$		t	$\Phi(t)$		t	$\Phi(t)$	
0,00	0,0000	399	1,00	0,6827	236	2,00	0,9545	51
0,05	0,0399	398	1,05	0,7063	224	2,05	0,9596	47
0,10	0,0797	395	1,10	0,7287	212	2,10	0,9643	41
0,15	0,1192	393	1,15	0,7499	200	2,15	0,9684	38
0,20	0,1585	389	1,20	0,7699	188	2,20	0,9722	34
0,25	0,1974	384	1,25	0,7887	177	2,25	0 9756	30
0,30	0,2358	379	1,30	0,8064	166	2,30	0,9786	26
0,35	0,2737	371	1,35	0,8230	155	2,35	0,9812	24
0,40	0,3108	365	1,40	0,8385	144	2,40	0,9836	21
0,45	0,3473	356	1,45	0,8529	135	2,45	0,9857	19
0,50	0,3829	348	1,50	0,8664	125	2,50	0,9876	16
0,55	0,4177	338	1,55	0,8789	115	2,55	0,9892	15
0,60	0,4515	328	1,60	0,8904	107	2,60	0,9907	13
0,65	0,4843	318	1,65	0,9011	98	2,65	0,9920	11
0,70	0,5161	306	1,70	0,9109	90	2,70	0,9931	9
0,75	0,5467	296	1,75	0,9199	82	2,75	0,9940	9
0,80	0,5763	284	1,80	0,9281	76	2,80	0,9949	7
0,85	0,6047	272	1,85	0,9357	69	2,85	0,9956	7
0,90	0,6319	260	1,90	0,9426	62	2,90	0,9963	5
0,95	0,6579	248	1,95	0,9488	57	2,95	0,9968	5
1,00	0,6827		2,00	0,9545		3,00	0,9973	

Таблица II

t	$\Phi(t)$	$1-\Phi(t)$	t	$\Phi(t)$	$1-\Phi(t)$
			2,878	0,996	0,004
1,960	0,95	0,05	2,968	0,997	0,003
2,054	0,96	0,04	3,090	0,998	0,002
2,170	0,97	0,03	3,291	0,999	0,001
2,326	0,98	0,02	3,481	0,9995	0,0005
2,576	0,99	0,01	3,891	0,9999	0,0001
2,612	0,991	0,009			
2,652	0,992	0,008	4,417		10^{-5}
2,697	0,993	0,007			
2,748	0,994	0,006	4,892		10^{-6}
2,807	0,995	0,005	5,327		10^{-7}

Пояснения см. на стр. 47, 53.

Лев Зимонович Румшиский.
Элементы теории вероятностей.

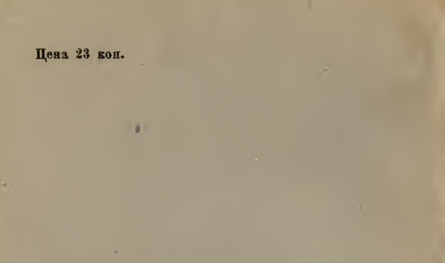
М., Физматгиз, 1963 г., 156 стр. с илл.

Редакторы *И. А. Брин* и *И. Е. Морозова.*
Техн. редактор *К. Ф. Брудно.*
Корректор *С. Н. Емельянова.*

Печать с матриц. Подписано к печати 31 I 1963 г. Бумага $84 \times 108^1/_{32}$. Физ. печ. л. 4,875. Условн. печ. л. 8,0. Уч.-изд. л. 7,58. Тираж 40 000 экз. Т-01536.
Цена книги 23 коп. Заказ № 100.

Государственное издательство
физико-математической литературы.
Москва, В-71, Ленинский проспект, 15.

Первая Образцовая типография
имени А. А. Жданова
Московского городского совнархоза.
Москва, Ж-54, Валовая, 28.



Цена 23 коп.